THE UNIVERSITY
of EDINBURGH

Intragenomic conflict under paternal genome elimination

Andrés G. de la Filia

Thesis submitted for the degree of
Doctor of Philosophy

School of Biological Sciences

The University of Edinburgh

2018

Table of contents

Thesis abstract

Genetic systems are extraordinarily variable across taxa, not only among large taxonomic groups but also between closely related species. The causes of this diversity are not fully understood, remaining a fundamental question in evolutionary biology. Genomic conflicts arising during reproduction might play an important role in shaping the striking diversity of reproductive strategies across life. Among these is paternal genome elimination (PGE), a form of haplodiploidy which has independently evolved several times in arthropods. Under PGE, males are diploid but transmit maternally-inherited chromosomes only. Moreover, in many forms of PGE, paternal chromosomes are silenced early in development. Due to the unequal inheritance patterns of PGE, mothers enjoy a transmission advantage through their sons, generating a strong scope for intragenomic conflict between paternally- and maternally-inherited genomes within males. However, empirical evidence for such conflict is lacking. In this thesis, I study whether paternal genomes can resist silencing and elimination in two PGE species: the citrus mealybug *Planoccocus citri* (Hemiptera: Pseudococcidae), where paternal chromosomes are silenced, and the human louse *Pediculus humanus* (Phthiraptera: Pediculidae), where PGE was recently discovered. I show that elimination of paternal chromosomes is leaky in both species and determine genome-wide patterns of parent-of-origin-specific expression to reveal the extent of paternal genome expression. Together, my findings offer a solid empirical ground to further explore the role of conflict in evolution of PGE and illustrate how this genetic system can impact different biological processes, such as hybrid incompatibilities and evolution of resistance.

Lay summary

In sexual reproduction, mothers and fathers tend to contribute one copy of half their genes, at random, to their sons and daughters. Sons and daughters are a mixture of their mother and father, and when they produce sperm or eggs, these will incorporate genes that they received from both their parents. This is the case in humans and almost every animal species on Earth—but not all.

In some insects, the genes that are transmitted by the males are always those that they received from their mothers. This strange mode of reproduction, called paternal genome elimination, generates an unfair situation for the genes that a male receives from his father, because they will not be transmitted. Moreover, the genes from the fathers are often deactivated in males, so only the genes from the mothers are expressed. Since the ultimate agenda of a gene is to pass on copies of itself to the next generation, paternal genes are expected to fight maternal genes within males so they can also be transmitted. This thesis explores if there is evidence for this fight between paternal and maternal genes in two insect species that have paternal genome elimination, the citrus mealybugs and the human louse, by studying if males can also transmit their fathers' genes in families of mealybugs and lice and if the fathers' genes are completely deactivated in males in these two species.

Acknowledgements

“And Frodo wouldn't have got far without Sam,
would he, dad?”

J. R. R. Tolkien, The Two Towers

I wouldn’t have got far either without the help and support of many people who have helped me carry this PhD until the very end. First and foremost, I am very grateful to my supervisor, Laura Ross, for introducing me into the mysterious and wonderful land of reproductive systems and for her constant support, guidance and encouragement. After four long years, I am still amazed by her scientific insight, imagination and ability to see the forest but not the trees, which are worth of the highest praise.

I am also obliged to my committee members, Darren Obbard and Konrad Lohse, for valuable discussions about my progress. I would also like to mention all the members of the Ross lab, most especially Dominik Laetsch for guiding me in my treacherous descent into the depths of bioinformatics. To someone who had barely open a terminal window before, his endless patience and advice were invaluable. Also, I am grateful to David Brown, who was instrumental in keeping our mealybug cultures alive, and all those who helped with the experiments presented here: Gyda Fenn-Moltu, Samuel Andrewes and Ross Cavalieri. I am very thankful to the Darwin Trust of Edinburgh for funding my PhD studies.

Hugest professional and above all personal thanks to Stevie Bain, the other half of the Pain and Panic partnership. We started this adventure side by side and have overcome together many pitfalls and disasters, but there were always laughs on the way. There are no words, in English or Spanish, for her infinite support, companionship and sense of humour. Looking back, it is a remarkable achievement that we have managed to keep each other relatively sane. She has really been my Sam—and I hope I have been hers.

There are many people who have greatly contributed to my enjoyment of these four years. They have either been climbing partners, travel companions, fellow moviegoers, flatmates, football enthusiasts, snacks providers, pub explorers, video game buddies, or a combination of those; above all, friends. In alphabetical order to avoid comparative aggravations (yes, Ed, I am thinking of you): Tom Booker, Liam Brierley, Eileen Butterfield, Luiz Carvalho, Gytis Dudas, Lisa Gecchele (you didn't get a paragraph but you can enjoy this parenthesis), Tom Godfrey, Ed Ivimey-Cook, Kat Keogan, Rasmus Lindberg, Nathan Medd, Amy Munro-Faure, Billy Palmer, Charlotte Repton, Cristina Rosique, Mafe Torres, Nora Villamil, Richard Whittet and many others.

For my New England days, I am very thankful to the Clark lab at the University of Massachusetts Amherst for hosting me and to Ben and Roxanna Normark for letting me stay in their beautiful home and burden their Scrabble games.

There are other people I would also like to thank. Laure, who shared part of this journey and was always supportive; Iván, loyal comrade and confidante; Carlos, Pablo, Ana and José for coming all the way here so we could head north and explore the country together; my friends from back home and previous adventures in foreign lands for making me feel like time has not passed when we reunite.

I also think I should have some words for all the mealybugs and lice (and their endosymbionts) that I have sacrificed in my experiments. As a biologist, one really feels the vertigo of evolutionary time thinking of a constant stream of genes making their way through uncountable perils over billions of years to end up crushed in an Eppendorf tube. In the words of a certain Rebel leader, many insects died to bring us this information.

Finally, I want to thank my parents, Paloma and Fernando, for having supported, stimulated and encouraged me for as long as I can remember and, especially, over these long years of studies and travels between countries (and, of course, for the reasonably satisfying alleles I inherited from them); my sister Irene, a doctor of another kind; and all members of my family, whom I have seen way less than I would have liked over the past years. Particularly I think of my uncle Emilio, who always showed great curiosity for the weird creatures I ended up studying despite my poor attempts to talk about my research in layman's terms. This thesis is dedicated to his memory.

Declaration

I declare that this thesis has been composed by myself and that it has not been submitted, in whole or in part, for any other degree or professional qualification. I confirm that the work contained herein is solely my own except where explicitly indicated below. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

Chapter 2. I carried out data collection and analysis. G. Fenn-Moltu conducted the majority of sex pheromone response trials with my assistance.

Chapter 3. I carried out data collection and analysis. D. Laetsch provided assistance with bioinformatics analysis and wrote the asa2 script, which was designed jointly by both of us.

Chapter 4. I carried out data collection and analysis. S. Andrewes (University of Massachusetts Amherst) carried out experimental crosses and S. Michaelides (University of Oxford) contributed to data collection. This chapter is presented as published.

Chapter 5. I carried out data collection and analysis. R. Cavalieri (University of Massachusetts Amherst) provided assistance with experimental crosses.

Appendix 1. I wrote the manuscript jointly with S. Bain and L. Ross. This appendix is presented as published.

All experiments presented in this thesis were designed by myself and L. Ross.

Andrés García de la Filia Molina

Chapter 1

General introduction

1.1. Intragenomic conflict and evolution of genetic systems

1.1.1. The diversity of sexual reproduction

Sexual reproduction, which is practically universal to eukaryotic organisms (Schurko *et al.* 2009), is extraordinarily diverse. A wide heterogeneity of genetic systems and sex determination mechanisms, some of which of remarkable complexity, have been described in eukaryotes (Normark 2003; Barrett 2010; Billiard *et al.* 2012; Bachtrog *et al.* 2014; Beukeboom & Perrin 2014; Ashman *et al.* 2014) (Table 1.1). This diversity is found not only between very distantly related eukaryotes, but also within kingdoms and lower taxonomic units—even between closely related species (Ashman *et al.* 2014). Why we observe so much variability in such a fundamental trait is a puzzling question, since genetic systems (i.e. the mode of organisation and transmission from parents to offspring of genetic material in a species) have profound implications in virtually every aspect of the biology of an organism. For example, microevolutionary processes (Charlesworth & Charlesworth 1995; Hedrick & Parker 1997; Parker & Hedrick 2000; Otto 2009), genome organization (Blackmon *et al.* 2016), offspring sex ratios (Hamilton 1967), inbreeding (Werren 1993; Henter 2003), sexual selection (Charnov 1979; Boulton *et al.* 2015), patterns of family relatedness (Hamilton 1964; Trivers & Hare 1976) and reproductive ecology (Appendix 1), among many other areas, have been shown to be majorly affected by different genetic systems. However, our understanding of the evolutionary forces and processes that shape the diversification of sexual reproduction is still limited, remaining a fundamental question in modern evolutionary biology.

Most eukaryotic organisms are diplodiploid: males and females (*sensu* Hurst & Hamilton 1992) each contribute a copy of a haploid genome to both their sons and daughters. Diplodiploidy is a symmetric genetic system: barring cytoplasmic genes or sex-specific chromosomes (Birky 1995; Bachtrog *et al.* 2014), each parent transmits half of their genetic material to their offspring, which are in turn equally likely to express and transmit alleles independently of the parent they inherited them from (Normark 2006). These symmetries in gene expression and transmission are the general rule in sexual reproduction (Wright 1931) and are ancestral to every sexual lineage with separate

sexes (Normark & Ross 2014), yet have given rise to many alternative genetic systems across the eukaryotic tree of life. These alternative genetic systems are often asymmetric, as they are characterised by differential genetic contributions of males and females to the offspring. For example, in some of these systems, females monopolise the parentage of one of the two sexes by parthenogenetically producing diploid daughters (thelytoky) or haploid sons (arrhenotoky, i.e. true haplodiploidy) (Normark 2014). In other systems, such as androgenesis, males are the sole genetic contributors to the offspring by excluding the maternal genome from the embryo (Schwander & Oldroyd 2016).

Table 1.1. A catalogue of the different genetic systems mentioned in this thesis, with representative examples of animal taxa where they are found.

Gonochoristic systems *(i.e. where individuals differentiate as one of at least two separate sexes)*		
Diplodiploidy Males and females transmit half of their genetic material to both sons and daughters, which express and transmit alleles independently of the parent they inherited them from	Most vertebrates and insects	Normark 2006
Haplodiploidy Males transmit maternally-inherited chromosomes only to their offspring, while females transmit both maternal and paternal alleles	Some insects, mites, rotifers	Normark 2003, de la Filia *et al.* 2015
Arrhenotokoky *(i.e. haplodiploidy sensu stricto)* Males develop from unfertilised eggs and are haploid, while females develop from fertilised eggs	Hymenoptera, Thysanoptera	Normark 2004
PGE *(a.k.a.* ***pseudoarrhenotoky****,* ***parahaploidy****)* Males and females develop from fertilised eggs, but males lose their paternally-inherited chromosomes early in development *(embryonic PGE)* or immediately before or during spermatogenesis *(germline PGE)*	Mealybugs, sciarid flies	Gardner and Ross, 2014
Parthenogenesis *(a.k.a.* ***thelytoky****)* Females produce diploid daughters from unfertilised eggs without paternal contribution	Bdelloid rotifers, *Timema* stick insects	Normark 2003
Androgenesis Males produce offspring that contain paternally-derived chromosomes only by hijacking eggs produced by females	*Corbicula* clams, *Bacillus* stick insects	Schwander & Oldroyd 2016
Non-gonochoristic systems *(i.e. where individuals do not differentiate as one sex)*		
Sequential hermaphrodites Individuals switch to the opposite sex during their lifetime	Many fish species (e.g. *Amphiprion percula*)	Ghiselin 1969, Avise & Mank 2009
Simultaneous hermaphrodites Individuals have both male and female reproductive organs simultaneously	Flatworms, *Caenorhabditis elegans*	Charnov 1979

These asymmetric patterns of gene inheritance generate conflicting consequences for males and females, which are directly translated into paternally- and maternally-inherited genes in the offspring. For example, under arrhenotoky maternally-derived alleles outnumber paternally-derived alleles by a proportion of 2:1 in the offspring (assuming equal sex ratios), since males only contain a haploid copy of the maternal genome (Kraaijeveld 2009). Moreover, an arrhenotokous mother guarantees full transmission of her genetic material through her sons, in addition to 50% (on average) through her daughters. In contrast, a father's genes can only be passed on to following generations through his daughters, who will incorporate 50% of their paternally-inherited alleles into their eggs. This conflict became apparent with the advent of the gene-centred view of evolution (Dawkins 1976) and the recognition of opposing evolutionary agendas in different genetic entities (Burt & Trivers 2006). Today, the notion of genomic conflict is central to genetics and evolutionary biology (Rice 2013). Genomic conflict offers a powerful framework to understand the evolutionary consequences of asymmetric genetic systems, but also lies at the heart of a growing body of evolutionary hypotheses aiming to explain the very emergence of such modes of reproduction.

1.1.2. Intragenomic conflict

Intragenomic conflict is a particular type of genomic conflict which arises when different entities within an individual genome have opposing evolutionary interests (Burt & Trivers 2006; Werren 2011; Gardner & Úbeda 2017). Interests of many genomic elements collide frequently, due to differences in transmission to the next generation or the action of divergent selective pressures: for example, nuclear versus cytoplasmic genes, such as heritable organelles and endosymbionts (Cosmides & Tooby 1981); autosomes versus sex chromosomes (Jaenike 2001); genomic parasites such as transposons and other mobile elements (Werren 2011) or supernumerary B chromosomes (Camacho *et al.* 2000) versus the rest of the genome; and paternally and maternally-derived alleles (Burt & Trivers 1998; Haig 2000). A wide range of biological phenomena have been attributed to intragenomic conflicts, and the consequences of such conflicts can lead to deleterious effects for the individuals (Crespi & Summers 2005; Úbeda & Wilkins 2008; Gardner & Úbeda 2017) but also to emergence of evolutionary novelty (Hurst 1992; Hurst & Werren 2001).

The most evident manifestation of intragenomic conflict occurs when a gene (*sensu* Gardner & Úbeda 2017) becomes able to enhance its own transmission to the following generation at the expense of other parties. The paradigmatic example of this intragenomic conflict is meiotic drive, whereby a gene (*sensu* Gardner & Úbeda 2017) hijacks gametogenesis to increase the number of its own copies that will be transmitted to the offspring (Lyttle 1993; Lindholm *et al.* 2016; Lenormand *et al.* 2016). Meiotic drive can act through non-random segregation during meiosis—e.g. driving centromeres in female monkeyflowers (Fishman & Willis 2005)—or by eliminating competing gametes—e.g. the SD gene complex in *Drosophila melanogaster* (Larracuente & Presgraves 2012), t-haplotypes in *Mus* (Lyon 2003), spore killers in fungi (Raju 1994). On the other hand, intragenomic conflict can emerge over the parental origin of genes within an individual, rather than their destination (Gardner & Úbeda 2017). The evolution of genomic imprinting, whereby genes display different behaviours (e.g. expression patterns) depending on the parent they were inherited from (Crouse 1960; Reik & Walter 2001; Ferguson-Smith 2011), has been framed as a consequence of intragenomic conflict between maternally- and paternally-inherited genes (Moore & Haig 1991; Burt & Trivers 1998; Haig 2000). The classic example of this kinship theory of genomic imprinting is the disagreement between parental alleles over allocation of maternal resources to the fetus in the placenta of eutherian mammals, with paternally-imprinted alleles promoting fetal growth and maternally-imprinted alleles restricting it (Moore & Haig 1991; Fowden *et al.* 2006).

Intragenomic conflicts are thought to have an evolutionary impact in different aspects of sexual reproduction, from emergence of hybrid incompatibilities (Frank 1991; Hurst & Pomiankowski 1991), sex allocation decisions (Shuker *et al.* 2009), sex determination (Werren & Beukeboom 1998) and sex chromosomes (Meiklejohn & Tao 2010) to evolution of sex itself (Hurst 1995). In particular, the idea that intragenomic conflicts may also shape the diversity of sexual reproduction by promoting the evolution of asymmetric genetic systems has attracted extensive interest. For example, parthenogenetic reproduction can arise as the outcome of an intragenomic conflict between nuclear and cytoplasmic genes, when maternally-transmitted endosymbionts take control of reproduction by eliminating the production of males (Werren *et al.* 2008; Stouthamer *et al.* 2010). Conflicts between parentally-inherited alleles have also been

suggested to underpin the evolution of asymmetric systems (Brown 1964; Bull 1979; Haig 1993; Herrick & Seger 1999; Normark 2006; Ross *et al.* 2010a; Gardner & Ross 2014). Under this view, asymmetric forms of reproduction such as haplodiploidy can evolve from a diplodiploid ancestral system when maternally- or paternally-inherited genomes take control of reproduction to maximise their transmission at the expense of their parental counterpart. This hypothesis is currently supported by several theoretical and verbal models with strong explanatory power, but is hindered by the lack of direct empirical validation. Some asymmetric modes of reproduction constitute promising study systems to explore the role of intragenomic conflict between parental alleles in an experimental context. One of such systems is paternal genome elimination, an extraordinarily bizarre genetic system that combines two major manifestations of intragenomic conflict—whole-genome meiotic drive with genomic imprinting.

1.2. Paternal genome elimination (PGE): an unusual genetic system

1.2.1. What is PGE?

Paternal genome elimination is an asymmetric genetic system that has independently evolved in at least seven arthropod lineages (Gardner & Ross 2014; Blackmon *et al.* 2015). Under PGE, males develop from fertilised eggs but only transmit maternally-inherited chromosomes to the offspring. Paternally-inherited chromosomes are excluded from sperm either by being destroyed prior to spermatogenesis or by being differentially segregated from the maternal set in an achiasmatic meiosis. Females, on the other hand, are normally diploid, exhibit meiotic recombination and transmit both maternally- and paternally-inherited alleles in a Mendelian fashion (Normark 2003; Burt & Trivers 2006) (Fig 1.1).

Since these transmission patterns are identical to arrhenotoky, PGE is considered a form of haplodiploidy and has also been referred to in the literature as parahaploidy (Hartl & Brown 1970) or pseudoarrhenotoky (Schulten 1985). The key difference with arrhenotoky is that PGE males are always diploid at the start of development, in contrast to arrhenokous males which develop parthenogenetically from

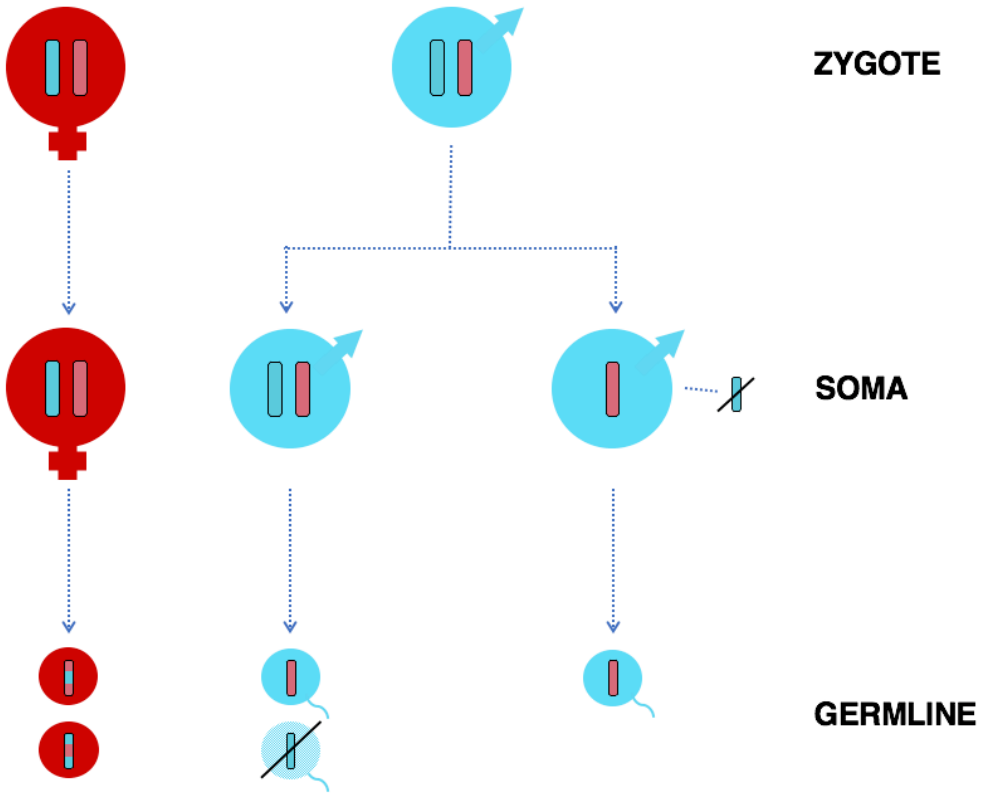


Figure 1.1. Haplodiploid-like gene transmission pattern under PGE. Both sexes develop from fertilised eggs. Maternally- and paternally-inherited haploid genomes are showed in pink and blue, respectively. Females are normally diploid, undergo recombination during meiosis and produce recombinant haploid eggs containing a mixture of the parental genomes. In males, paternal chromosomes are destroyed during spermatogenesis or eliminated earlier from the soma, so that only maternally-inherited chromosomes enter spermatogenesis. In both cases, only the maternal genome is incorporated into active sperm.

unfertilised eggs without a paternal contribution. In both systems, these asymmetric gene inheritance patterns confer a transmission advantage to females through sons: arrhenotokous and PGE males exclusively incorporate maternally-derived chromosomes into active sperm, so that all fertilised eggs (females under arrhenotoky, both sexes under PGE) will receive a complete copy of the male's maternal genome.

To date, PGE is estimated to be present in ~20,000 arthropod species (Gardner & Ross 2014). Clades where PGE has evolved include scale insects (Hemiptera: Neococcidea) (Nur 1980; Ross *et al.* 2010a), fungus gnats (Diptera: Sciaridae) (Goday & Esteban 2001) and gall midges (Diptera: Cecidomyiidae) (Burt & Trivers 2006). These

are remarkably species-rich lineages (Gullan & Cook 2007; Espírito-Santo & Fernandes 2007; Shin *et al.* 2013) with deep evolutionary roots, revealing that PGE is a successful reproductive strategy that persists over long evolutionary timescales. Additionally, PGE has been found in a bark beetle, the coffee bean borer *Hypothenemus hampei* (Coleoptera: Curculionidae) (Brun *et al.* 1995), and two louse species, the booklouse *Liposcelis* (Psocoptera: Liposcelididae) (Hodson *et al.* 2017) and the human body louse *Pediculus humanus humanus* (Phthiraptera: Pediculidae) (McMeniman & Barker 2005). Finally, two non-insect arthropod clades show PGE: springtails (Collembola: Symphypleona) and at least three families of predatory mites (Acari: Mesostigmata) (Nelson-Rees *et al.* 1980; Burt & Trivers 2006). This diverse range of organisms are often characterised by high levels of inbreeding, female-biased sex ratios, maternally-inherited endosymbionts, unusual sperm structures and highly modified meiosis (Normark 2004; Burt & Trivers 2006; Gardner & Ross 2014; Ross & Normark 2015), which suggests that PGE could also be present in additional arthropod groups which share these characteristics.

Exclusion of paternal chromosomes from the germline is a common feature to all independent origins of PGE, but the precise timing of their elimination shows variability among groups (Herrick & Seger 1999; Gardner & Ross 2014). Depending on when paternal chromosomes are destroyed, two forms of PGE can be broadly recognised: embryonic and germline PGE. In embryonic PGE, which is present in mites and armoured scale insects, paternal chromosomes are destroyed during early development and males remain completely haploid for the rest of their lives. In germline PGE, paternal chromosomes are retained in the soma and elimination is delayed until spermatogenesis. However, different germline PGE taxa vary in the extent to which the paternal genome is retained and expressed in somatic cells. In springtails and sciarid flies, one or two chromosomes are eliminated during development, so that only a fraction of the paternal genome is retained in males. In other groups, such as soft scale insects, the coffee borer beetle and booklice, all paternal chromosomes are present in somatic cells, yet they are transcriptionally silenced in development and remain inactive until their elimination.

This remarkable variability in timing of loss and degree of expression of paternal chromosomes among PGE is still poorly understood. More strikingly, even closely

related lineages show diversity in the form of PGE they display. This is the case of the scale insects, by far the best studied of the PGE lineages (Hughes-Schrader 1948; Brown 1963; Brown 1967; Nur 1980; Ross *et al.* 2010a) and currently the most promising group to study the evolutionary dynamics of this unusual genetic system.

1.2.2. PGE in scale insects

Scale insects are a group of external plant parasites that feed on the host phloem and have developed strong adaptations to this lifestyle (Miller & Kosztarab 1979; Gullan & Kosztarab 1997). Sexual dimorphism is extreme: adult females are wingless, typically sedentary and covered by a waxy cuticle or real scales, while males are often smaller, winged and have a much shorter adult lifespan. Both sexes are undistinguishable during early larval instars and start to differ in late instars, when males stop feeding and pupate while females continue growing and retain paedomorphic morphology.

PGE evolved once in scale insects (Yokogawa & Yahara 2009), most likely from a XX-XO diplodiploid system which is believed to be the ancestral mode of reproduction in this group (Nur 1980). In addition to PGE, the diverse array of alternative genetic systems that have evolved in this group include arrhenotoky, hermaphroditism and parthenogenesis (Ross *et al.* 2010a). The instability in genetic systems found in scale insects is also reflected by the diversity of forms of PGE that have been described within the group (Fig. 1.2). The most basal form of PGE in scale insects is known as lecanoid PGE (Hughes-Schrader 1948). Lecanoid PGE is of the germline type, with silencing of paternal chromosomes in early development and elimination during spermatogenesis. Lecanoid PGE is found in three scale insect families, including mealybugs (Pseudococcidae), the group where PGE was first discovered (Schrader 1921; Schrader 1935). A second form of PGE, known as Comstockiella, has repeatedly evolved from the lecanoid system (Brown 1963; Nur 1980). Comstockiella PGE is very similar to lecanoid PGE, but differs in that a small number of paternal chromosomes are destroyed prior to spermatogenesis. The third type, diaspidid PGE, is embryonic: paternal chromosomes are eliminated, instead of silenced, during early development. Diaspidid PGE is found in armored scale insects (Diaspididae), and its evolution is nested within the Comstockiella system (Nur 1980). Transitions between forms of PGE in scale insects are common and can happen at short evolutionary timescales (Herrick & Seger 1999; Ross *et al.* 2010a).

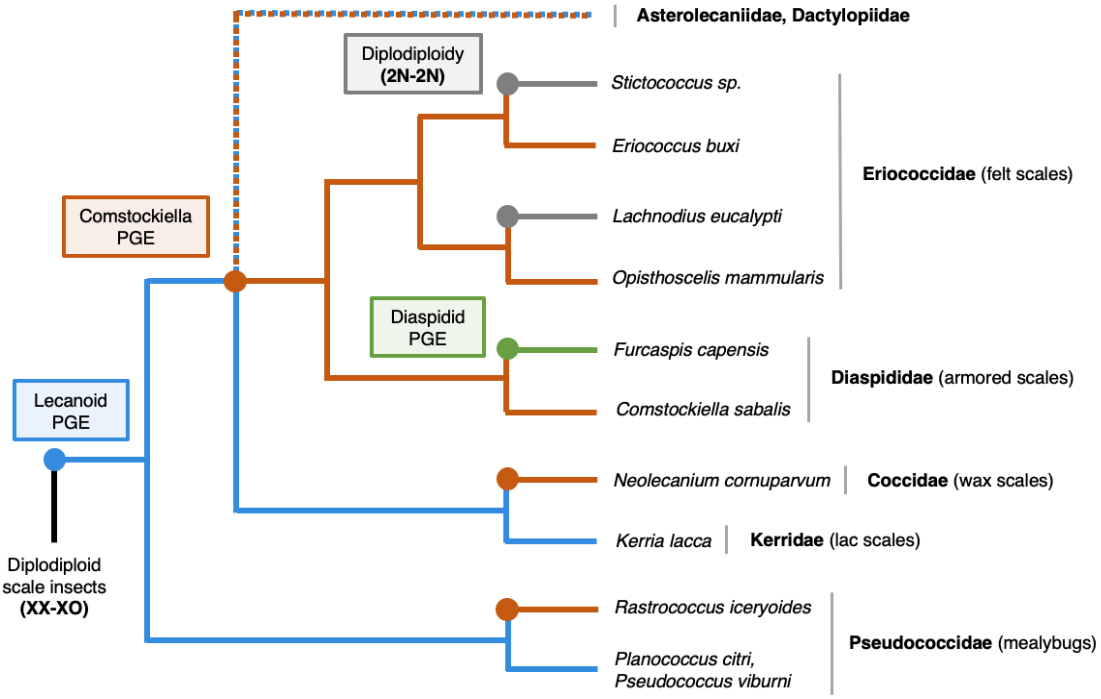


Figure 1.2. Partial scale insect phylogeny, reconstructed from Gullan & Cook 2007, Gavrilov 2007 and Ross *et al.* 2010a (branches not to scale). All families with PGE are represented, with different forms of PGE or diplodiploidy indicated by coloured lines and transitions highlighted by circles. Representative species from these PGE families are shown.

More strikingly, in spite of the long evolutionary history of PGE in scale insects, this instability has also led to loss of PGE in at least two scale insect genera: *Stictococcus* and *Lachnodius* (Ericoccidae), which have independently reverted to diplodiploidy (Normark 2003). The reason behind these frequent transitions seems not lie within ecological factors or life histories, which are remarkably similar across all scale insects. A solid body of theoretical work based on intragenomic conflict has been built to explain the evolution of this system and its different manifestations.

1.3. The evolution of PGE: an overview of the conflict hypothesis

1.3.1. Theoretical models on the emergence and diversification of PGE

Under PGE, females enjoy a transmission advantage through their sons, as they will always transmit maternal alleles to the following generation. The evolution of PGE

was first linked to this maternal transmission advantage in a seminal paper by Brown (1964) focusing on the selective pressures leading to male haploidy in scale insects (functional, in lecanoid coccids, or actual, in diaspidids). In his model, Brown recognised the similarities between PGE and meiotic drive and proposed that maternal genes leading to male haploidy could spread in a population, since their transmission is enhanced through haploid males compared to diploid males even if male fitness decreases due to reduction in haploidy. Later, in a more general model, Bull (1979) explicitly framed PGE under intragenomic conflict and extended Brown's model to a wider set of conditions upon which male haploidy can arise under the control of maternal genes. Importantly, Bull's model made two vital advances. First, he implicitly introduced the notion of antagonistic coevolution between parentally-inherited alleles by also considering a selective advantage for paternal genes to restore diploidy after PGE has arisen. Second, he generalised his model to other systems where transmission of maternal genomes is increased, such as arrhenotoky and thelytoky, showing that maternal-paternal conflicts over transmission could be a general force driving the evolution of different non-diploid systems. An important limitation of these early models is their dependence on a non-male heterogametic sex determination system for evolution of PGE, since all sperm would bear an X chromosome and produce all-female offspring, leading to population extinction. Therefore, Haig (1993) proposed a new model based on the ancestral XX-XO system in scale insects leading to the fixation of PGE in three steps: i) initial evolution of X-chromosome drive in males, resulting in female-biased (but not female-only) sex ratios ii) linkage of maternal autosomal genes to the driving X to extend transmission advantage to the whole maternal genome and iii) shift in sex determination mechanism into maternal control, allowing for fixation of PGE. More recently, Gardner & Ross (2014) explored the invasion of parental genome drive in the context of the reproductive ecology of PGE taxa and demonstrated how PGE, and not other forms of genome elimination, can become fixed in a population characterised by high levels of inbreeding and male heterogamety.

Due to their prevalence in the literature, these theoretical approaches on the evolution of PGE from diplodiplody were founded on scale insects. The nested evolution of different forms of PGE within this group, however, remains an intriguing phenomenon that cannot be satisfactorily explained with these models. Building on the initial maternal

transmission advantage that resulted in emergence of PGE, Herrick & Seger (1999) interpreted the successive transitions between the lecanoid, Comstockiella and diaspidid forms of PGE as different stages of a coevolutionary arms race between paternal and maternal genomes. They suggested that the most basal form of PGE, characterised by germline elimination of paternal chromosomes only, would be ripe for evolution of paternal suppressors of maternal drive to restore a fair spermatogenesis. These anti-PGE adaptations allowing all or a fraction of the paternal genome to be incorporated into sperm would rapidly spread, even if their success was limited, so they would elicit maternal counter-responses to maintain transmission advantage. Suppressing expression of paternal alleles by inducing heterochromatinization in early development would be an effective way to control breakouts, leading to evolution of lecanoid PGE. However, decondensation of paternal chromosomes would allow reactivation of anti-PGE adaptations, so elimination prior to their entrance in spermatogenesis, as in the Comstockiella system, would have been a subsequent maternal move. Eventually, complete obliteration of paternal resistance would be achieved by bringing elimination of paternal genes forward during male development, destroying paternal chromosomes during male development and giving rise to the diaspidid system.

1.3.2. Current evidence for antagonistic coevolution between parental alleles

Herrick and Sager's model is currently accepted as the most plausible explanation for the evolutionary dynamics of PGE (Burt & Trivers 2006; Shuker *et al.* 2009; Ross *et al.* 2010a). Indeed, the transitions suggested in their model are supported by a wealth of observations. Their model states that heterochromatinization of paternal chromosomes is a maternal adaptation to prevent expression of paternal alleles, explaining the emergence of lecanoid PGE. In mealybugs, heterochromatinization takes place at the precise time zygotic transcription is initiated (Sabour 1972), and the presence of the maternal genome is necessary to maintain their heterochromatic status (Nur 1962b; Chandra 1963; Brown & Nur 1964).

Their lecanoid-Comstockiella transition is based upon the assumption than paternal chromosomes can become decondensed, thus regaining genetic activity, to escape elimination. Many observations support this hypothesis. First, radiation-induced decondensation of paternal chromosomes leads to failure to discriminate paternal from

maternal homologs during meiosis in mealybugs, resulting in diploid sperm (Nur 1970). Consistent with this finding, in the mealybug *Pseudococcus viburni* paternally-inherited supernumerary B chromosomes that are heterochromatinized in somatic tissues become decondensed during late prophase I and ultimately segregate with the maternal chromosomes (Nur 1962a). Second, the heterochromatic status of paternal chromosomes is naturally reversed in certain tissues of several mealybug species (Nur 1966; Nur 1967; Nur 1990). Outstandingly, one of the tissues where this reversal occurs are somatic parts of the testes during spermatogenesis, which supports the idea that paternal alleles play a role during spermatogenesis in lecanoid PGE (Nelson-Rees 1962). Finally, male offspring of hybrid crosses between mealybug species are often inviable (Nur & Chandra 1963; Rotundo & Tremblay 1982; Tranfaglia & Tremblay 1982; Kol-Maimon *et al.* 2014a), although complete silencing of paternal chromosomes should protect hybrid males from deleterious interactions between parental genomes. Additionally, in Comstockiella systems, the number and identity of chromosomes that are eliminated premeiotically shows remarkable variability not only between species, but also within species and even within individuals (Brown 1967; Kitchin 1975; Ross *et al.* 2010a), which could be suggestive of different intensities of paternal responses.

The phylogenetic distribution of transitions between forms of PGE within the Neococcidae tree also offers support to the arms race hypothesis. Lecanoid to Comstockiella transitions, and their reverse, have been frequent over a long evolutionary period in at least one scale insect family, Eriococcidae (Brown 1967; Gavrilov 2007). Comstockiella to diaspidid PGE shifts have independently occurred a minimum of four times within the Diaspididae family (Herrick & Seger 1999), but reversals from diaspidid to Comstockiella have not been found, supporting the view that paternal genome resistance is ultimately defeated with embryonic elimination. Finally, both known independent reversions from PGE to diplodiploidy have arisen from lecanoid (*Stictococcus*) and Comstockiella (*Lachnodius*) systems (Gavrilov 2007), where paternal resistance can still manifest itself and gain victory. In both genera, sex determination differs from the ancestral XX-XO of other diplodiploid scale insects as they lack a differentiated sex chromosome, which is consistent with the shift in sex determination following evolution of PGE predicted by Haig (1993). Although the involvement of endosymbiotic bacteria in the loss of PGE in *Stictococcus* has also been suggested

(Normark 2004), *Lachnodius* is identical in their ecology to closely related lecanoid genera, which leaves paternal resistance as the most plausible explanation for their reversal to diplodiploidy.

The parental genome arms race hypothesis has a great explanatory power in placing the origin and diversity of PGE into a likely evolutionary scenario. However, it has two important limitations. First, very few of its hypotheses and predictions have been empirically tested. Direct proof of paternal chromosomes escaping germline elimination or overcoming heterochromatization to regain genetic activity is lacking, and the precise maternal adaptations to exert control of the paternal set are yet to be revealed. Second, only a minimal fraction of the species exhibiting PGE has been sampled so far, so our current understanding of PGE is limited to a minority of well described taxa. Only a few intermediates between forms of PGE where conflict can become most apparent (for example, the unstable Comstockiella system in the Eriocccidae family) have been preliminarily explored, while taxa in incipient stages of PGE prior to fixation of maternal genome drive or where paternal genomes are successfully counteracting their elimination remain uncharacterised.

In this thesis, I aim to bridge this existing gap between theory and evidence on the role of intragenomic conflict in the evolution of PGE. To do so, I follow two routes: i) I explicitly test predictions of the arms race hypothesis in a well-establish PGE model, *Planococcus citri*, and ii) I fully describe the genetic system of a species which could constitute an intermediate step in evolution of PGE, *Pediculus humanus* (Fig. 1.3A).

1.4. Study systems

1.4.1. The citrus mealybug Planoccocus citri*: a model organism for PGE*

The citrus mealybug *Planococcus citri* (Hemiptera: Pseudococcidae) is arguably the PGE species that has been more extensively studied in the context of this genetic system since its early adoption as an insect model system for genomic imprinting and epigenetics (Brown & Nur, 1964; Bongiorni & Prantera, 2003; Khosla *et al.*, 2006; Ross

et al., 2010a; Prantera & Bongiorni, 2012). *P. citri* is a widespread agricultural pest that feeds primarily on the phloem of citrus plants, but is also able to live on a wide variety of host species (Franco *et al.* 2009). As all scale insects, both sexes display extreme sexual dimorphisms in late stages. The life cycle of the citrus mealybug spans 30 days, with males and females becoming differentiated after the second larval instar. Males stop feeding after this stage and undergo two additional prepupal and pupal instars before reaching adulthood, while females continue growing. Males become sexually mature after the first 24h of adulthood and only live up to 2-5 days. *P. citri* females are obligatorily sexual and can only lay eggs after mating (Da Silva *et al.* 2010). Their adult lifespan is longer, being able to live >60 days if unmated. *P. citri* forms an intricate tripartite

A

B

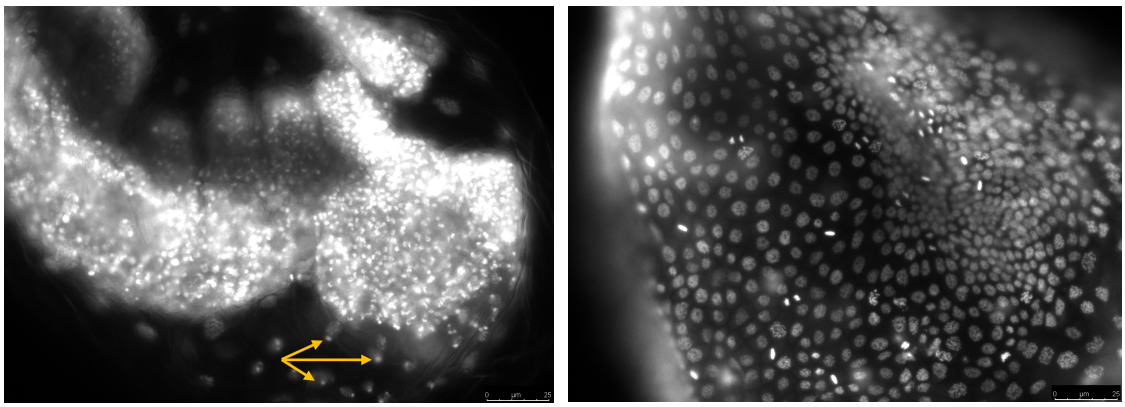


Figure 1.3. Study species **(A)** Left**,** the citrus mealybug *Planococcus citri*. A female (left, globular) and a male (right, winged) are shown during mating. Right, the human louse *Pediculus humanus.* An adult head louse female is shown next to an egg laid on a human hair **(B)** DAPI-stained *P. citri* embryos. Left, a male embryo exhibit conspicuous chromocentres within the cell nuclei (see examples marked with orange arrows). Right, a female embryo.

symbiotic system with two bacteria which provide essential amino acids and other nutrients absent from plant phloem: the primary endosymbiont is a β-proteobacterium, *Tremblaya princeps*, which in turn hosts the secondary endosymbiont, the γ-proteobacterium *Moranella endobia* (Von Dohlen *et al.* 2001; Husnik *et al.* 2013). *P. citri* lacks sex chromosomes, and the sex determination mechanism remains unknown (Hughes-Schrader 1948).

Although the initial discovery of somatic condensation of chromosomes in males was reported in a close relative (Schrader 1921), most of the following studies that led to characterization of the lecanoid PGE system were focused on *P. citri.* In males, paternal chromosomes are heterochromatinized at the onset of gene expression in the embryo (Sabour 1972). In early blastula stage, a heterochromatinization cascade proceeds from the tip of male embryos to condense their paternal genomes into tightly packed bodies known as chromocentres (Bongiorni *et al.* 2001) (Fig. 1.3B). From that point, the heterochromatic state of paternal chromosomes is faithfully transmitted to daughter cells during mitosis. Although the paternal origin and genetic inactivity of heterochromatic chromosomes was readily hypothesised (Schrader & Hughes-Schrader 1931; Hughes-Schrader 1948), conclusive empirical demonstration was provided in an elegant irradiation experiment (Brown & Nelson-Rees 1961). Mealybug chromosomes are holocentric: they lack a localised centromere so that kinetochore function is distributed along their entire length (Schrader 1935; Wrensch *et al.* 1994). Hence, chromosome fragments caused by irradiation are faithfully transmitted to the offspring (Hughes-Schrader & Ris 1941). In their study, Brown & Nelson-Rees (1961) demonstrated that viability of male and female offspring of irradiated mothers was similarly affected by deleterious effects of X-ray damage in the maternally-inherited chromosomes, which were euchromatic in both sexes. However, when fathers were irradiated, only daughters were affected, while sons were viable and showed the fragmented chromosomes in a heterochromatic state. Later, autoradiography essays showed that RNA synthesis was blocked from the heterochromatic chromosomes (Berlowitz 1965) and studies of inheritance of genetic markers confirmed that paternally-inherited traits were not expressed in males (Brown & Nur 1964; Brown & Wiegmann 1969; Brown 1972). However, during late larval instars, paternal chromosomes appear to lose their heterochromatic state in some or all cells from different tissues, including

oenocytes, skeletal muscle, intestinal tract, Malpighian tubules and cyst wall of testes (Nur 1966; Nur 1967). Loss of heterochromatization in parts of the testes is consistent with reports of elevated levels of paternal irradiation causing sterility in male progeny (Nelson-Rees 1962), which suggests an involvement of the paternal genome in spermatogenesis that remains unexplored.

As in all germline PGE taxa, elimination of paternal chromosomes is delayed until spermatogenesis, which has been meticulously described in mealybugs (Hughes-Schrader 1948; Bongiorni *et al.* 2004; Bongiorni *et al.* 2009). Spermatogenesis is initiated by spermatogonial precursor cells located at the tip of the testis, which undergo a series of four mitotic divisions to produce a cluster of 16 primary spermatogonia that simultaneously enter meiosis. Meiosis is inverted in mealybugs: sister chromatids separate during the first meiotic division and reduction of ploidy is delayed until meiosis II (Chandra 1962; Viera *et al.* 2008) (Fig. 1.4). After the equational meiosis I, segregation of parental chromatids takes place in anaphase II, which involves a monopolar spindle that only interacts with the euchromatic maternal set while the heterochromatic paternal complement lags behind. The result of this process is a 64-cell cyst in which only half of the spermatids—those carrying maternal chromosomes—progress to complete sperm maturation and the remaining 32 spermatid nuclei containing the paternal sets degenerate in situ (Bongiorni *et al.* 2004).

To date, limited progress has been made in identifying the signals that differentially tag maternal and paternal chromosomes in *P. citri*. In mealybugs, the role of epigenetic marks such as DNA methylation (Bongiorni *et al.* 1999; Bongiorni *et al.* 2009) and histone modifications (Khosla *et al.* 2006; Prantera & Bongiorni 2012) has been investigated mostly in relation to their interactions with heterochromatin-associated proteins underlying somatic inactivation of paternal chromosomes, such as HP-1 and HP-2 homologs (Bongiorni *et al.* 2001; Bongiorni *et al.* 2007; Volpi *et al.* 2007). The prime candidate signal to distinguish parental chromosomes in somatic cells are levels of DNA methylation (Bongiorni *et al.* 1999), although results are inconclusive. Comparatively, less is known about how these signals are established during spermatogenesis or how they mediate differential segregation of spermatids during the second meiotic division. Bongiorni *et al.* (2009) showed that both HP-1 and HP-2 and several histone modifications are present during all stages of spermatogenesis and undergo rapid

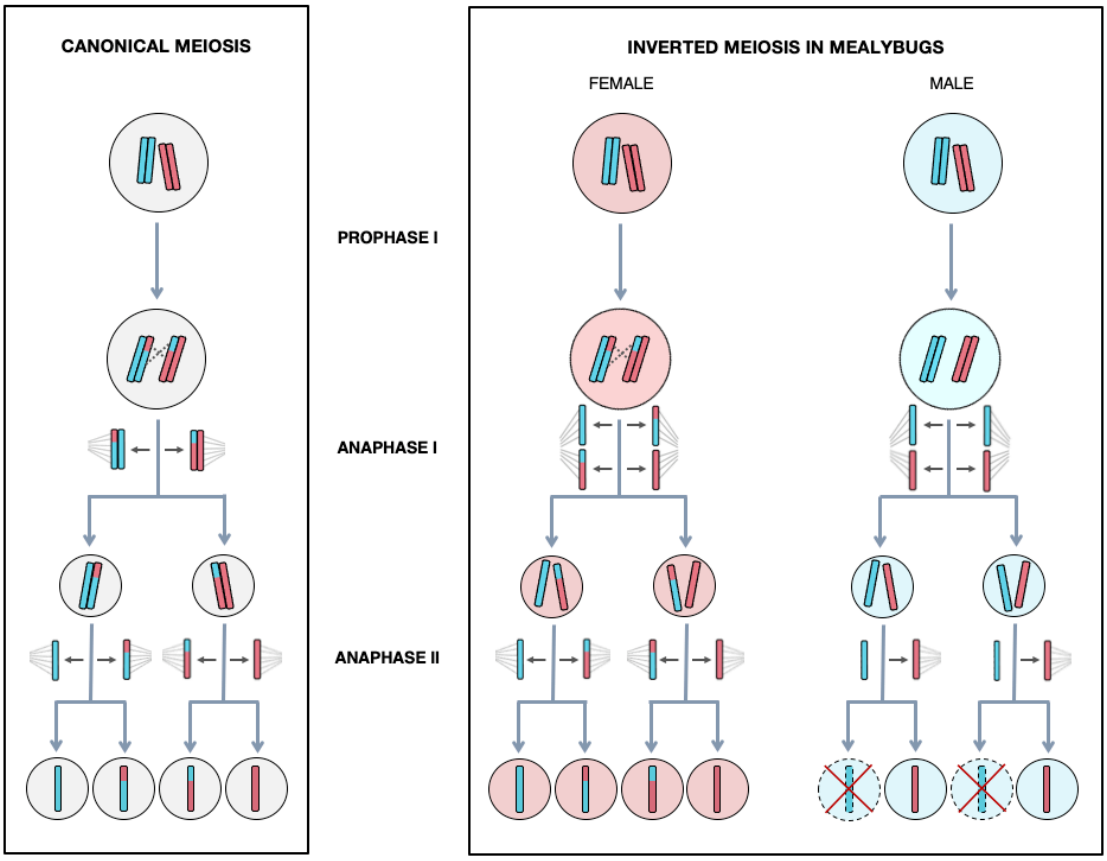


Figure 1.4. Inverted meiosis in mealybugs. In canonical meiosis (left), following recombination, homologous chromosomes separate in the first meiotic division (reductional) and sister chromatids segregate to different poles in meiosis II (equational). In inverted meiosis (right), the order of events is reversed: separation of sister chromatids takes place in meiosis I (equational) and homologous non-sister chromatids segregate to different poles in the meiosis II (reductional). In mealybug males, there is no recombination and differential segregation of paternal and maternal chromatids in each daughter cell during anaphase II involves a monopolar spindle that only interacts with maternal chromatids, while their paternal homologs lag behind.

redistributions along this process. In contrast, these marks cannot be detected at any stage during oogenesis, which suggests a key role in maintaining paternal status until segregation of paternal and maternal chromosomes in meiosis II. However, at present it is still difficult to deduce a role of these signals in primary and secondary spermatocytes.

1.4.2. The human louse Pediculus humanus*: an intermediate stage in evolution of PGE?*

Pediculus humanus (Phthiraptera: Pediculidae) is a cosmopolitan blood-sucking ectoparasite of humans. Traditionally, two subspecies of human lice (now considered ecotypes) have been recognised (Durden & Musser 1994): head lice (*Pediculus humanus capitis*), which feed on human scalp, and the body louse (*Pediculus humanus humanus*), which feed on skin and lay eggs on clothes. Head and body lice have identical life cycles: eggs hatch ~7 days after oviposition and males and females reach adulthood after 3 larval instars (~10 days) (Veracx & Raoult 2012). Both lice house an obligate nutritional endosymbiont, the γ-proteobacterium *Riesia pediculicola* (Sasaki-Fukatsu *et al.* 2006). Body lice also harbour several epidemic pathogenic bacteria, such as *Bartonella quintana* (trench fever), *Borrelia recurrentis* (relapsing fever) and *Rickettsia prowazekii* (epidemic typhus) (Raoult & Roux 1999).

PGE was recently reported in the human body louse, where it appears to be polymorphic (McMeniman & Barker 2005). While some males were found to transmit maternal alleles exclusively, others can transmit both maternal and paternal alleles in a Mendelian fashion. Two alternative explanations can account for this polymorphism: either PGE is still arising and the evolutionary transition from diplodiploidy has not been completed or PGE was already established and a germ-line elimination suppressor evolved recently. However, no follow-up studies have been published since the original report and several fundamental questions remain unexplored. There is no information on allele transmission patterns in head louse or to what extent paternal genomes are expressed in head or body lice.

1.5. Thesis aims and structure

This thesis is equally divided between empirical work on mealybugs and human lice. The main goal of this thesis is to address two biological questions that are fundamental in increasing our understanding of PGE and its evolution. First, this thesis explores whether there is an evolutionary arms race between maternal and paternal alleles under PGE, as predicted in the evolutionary models discussed above, by testing for evidence of conflict and coevolution between parental genomes in a well-established and thoroughly described PGE model system. These experiments are presented in the first two data chapters, focused on *P. citri* and the closely related *P. ficus*, where I

explicitly test whether there is evidence for a current or historical arms race between paternal and maternal alleles. In **Chapter 2**, I investigate if males can transmit paternal genotypes and heritable phenotypes to their offspring both at intraspecific and interspecific scales using a combination of microsatellite markers and sex pheromone response assays. In **Chapter 3**, I study genome-wide patterns of allele-specific gene expression in hybrid male mealybugs to determine whether silencing of paternal chromosomes is complete in both somatic and reproductive tissues, which are expected to be a hotspot for expression of paternal resistance under the arms race hypothesis.

The second question addressed by this thesis is to fully characterise gene transmission and expression patterns in human lice and thus determine, as suggested by available evidence, whether *P. humanus* represents a evolutionary intermediate between PGE and diplodiploidy. In **Chapter 4**, I use microsatellite markers to confirm the occurrence of polymorphic PGE in the body louse and determine if head lice exhibit the same mode of inheritance. In **Chapter 5**, I apply the allele-specific expression analysis pipeline employed in Chapter 3 to determine whether paternal genomes are expressed in lice. Finally, a recapitulation of main findings and suggestions of future research directions is presented in **Chapter 6.**

Chapter 2

Assessment of an intragenomic arms race under paternal genome elimination in mealybugs

2.1. Summary

Genomic conflicts arising during reproduction might play an important role in shaping the striking diversity of reproductive strategies across life. Among these is paternal genome elimination (PGE), a form of haplodiploidy which has independently evolved several times in arthropods. PGE males are diploid but transmit maternally-inherited chromosomes only, while paternal homologues are excluded from sperm. Mothers thereby effectively monopolize the parentage of sons, at the cost of the father's reproductive success. This creates striking conflict between the sexes that could result in a coevolutionary arms race between paternal and maternal genomes over gene transmission, yet empirical evidence that such an arms race indeed takes place under PGE is scarce. This study addresses this by testing if PGE is complete when paternal genotypes are exposed to divergent maternal backgrounds in intraspecific and hybrid crosses of the citrus mealybug, *Planococcus citri*, and the closely related *P. ficus*. We determined whether males can transmit genetic information through their sons by tracking inheritance of two traits in a three-generation pedigree: microsatellite markers and sex-specific pheromone preferences. Our results suggest leakages of single paternal chromosomes through males occurring at a low frequency, but we find no evidence for transmission of paternal pheromone preferences from fathers to sons. The absence of differences between hybrid and intraspecific crosses in leakage rate of paternal alleles suggests that a coevolutionary arms race cannot be demonstrated on this evolutionary timescale, but we conclude that there is scope for intragenomic conflict between parental genomes in mealybugs. Finally, we discuss how these paternal escapes can occur and what these findings may reveal about the evolutionary dynamics of this bizarre genetic system.

2.2. Introduction

Sexual reproduction is extremely variable, a result of the extraordinary diversity of genetic and reproductive systems that have evolved across the tree of life (Bachtrog *et al.* 2014). The sources of this variability remain elusive, so understanding which forces and processes drive transitions between genetic systems and the emergence of complex modes of reproduction is an important challenge for modern evolutionary biologists. One evolutionary force commonly invoked is intragenomic conflict (Burt & Trivers 2006; Ross *et al.* 2010a; Werren 2011; Normark & Ross 2014; Úbeda *et al.* 2015; Gardner & Úbeda 2017). Such conflict occurs when different genetic entities that coexist within individuals (e.g. nuclear versus cytoplasmic genes, autosomes versus sex chromosomes, mobile elements) disagree over transmission to following generations (Gardner & Úbeda 2017). In sexually reproducing eukaryotes, an important potential source of intragenomic conflict is the parental origin of the haploid copies that make up a diploid genome, as they are inherited from two different individuals with an evolutionary interest in maximising the transmission of their own genes (Normark & Ross 2014). Many alternative genetic systems emerge when mothers or fathers gain a transmission advantage by enhancing the transmission of the copies they transmit to the offspring at the expense of their partners': for example, arrhenotoky (i.e. true haplodiploidy), under which mothers monopolize parentage of sons, or androgenesis, where fathers are the sole contributors of genetic material to both offspring sexes (Normark 2006; Schwander & Oldroyd 2016). These systems are dramatic manifestations of intragenomic conflict, as Mendelian laws of fair inheritance are thwarted and genes undergo different fates depending on the sex of the individual they find themselves in.

One of the genetic systems where such conflict is particularly apparent is paternal genome elimination (PGE). PGE is a form of haplodiploid reproduction in which males develop from fertilized eggs (in contrast to arrhenotoky), but eventually lose their paternally-inherited chromosomes and only transmit the maternal homologs to the offspring (Normark 2003; Burt & Trivers 2006). PGE has a rich evolutionary history: it has independently evolved at least six times in insects and once in mites (Burt & Trivers 2006; Gardner & Ross 2014; de la Filia *et al.* 2015; Blackmon *et al.* 2015). Although males of all species with PGE lose their paternal chromosomes, timing of loss varies between taxa. In some groups, paternal chromosomes are lost early in development (embryonic PGE); in others, males remain (mostly or completely) somatically diploid and

elimination of paternal chromosomes is delayed until spermatogenesis, when they fail to be incorporated into active sperm (germline PGE). Moreover, some germline PGE taxa shut down expression of paternal chromosomes, which are highly condensed and therefore transcriptionally inactivated (Gardner & Ross 2014).

When considering transmission patterns of genes under PGE, it is clear why it leads to intragenomic conflict between maternal and paternal genomes: maternally-inherited alleles enjoy a transmission advantage through sons at the expense of paternally-inherited alleles, directly reflecting a conflict between male and female partners in which the latter have gained the upper hand (Brown 1964; Bull 1979; Haig 1993). Such conflict is likely to unchain an evolutionary arms race between both sexes and, consequently, maternally- and paternally-inherited alleles during spermatogenesis (Herrick & Seger 1999; Ross *et al.* 2010a). Once PGE has arisen, there is strong selection on males to evolve adaptations that will allow (all or a fraction of) their alleles to escape elimination when in sons. However, the success of these paternal adaptations is predicted to be short lived, as they will trigger the evolution of maternal responses to override paternal resistance and maintain complete transmission advantage of maternally-inherited alleles (Herrick & Seger 1999). Since germline PGE is a type of whole-genome meiotic drive in which the entire maternal chromosomal complement drives, the dynamics of this arms race in this system are similar to other drive-suppression systems (Burt & Trivers 2006; Lenormand *et al.* 2016; Lindholm *et al.* 2016). Examples of drive-suppression evolution include sex-linked alleles (Tao *et al.* 2007; Phadnis & Orr 2009), autosomal haplotypes (Schimenti 2000; Larracuente & Presgraves 2012), centromeric elements (Fishman & Willis 2005; Chmátal *et al.* 2014) and supernumerary chromosomes (Camacho *et al.* 2000). In these systems, when one of these genetic entities drives (i.e. develops the ability to manipulate meiotic processes to increase its presence in gametes at the expense of the rest of the genome), suppressors emerge to restore transmission symmetry.

To date, no empirical validation in support of these evolutionary scenarios is available. There is very scarce evidence of paternal escapes under PGE, which have only been explicitly shown in a PGE species, the human louse *Pediculus humanus* (Phthiraptera: Pediculidae) (McMeniman & Barker 2005). Direct empirical evidence for a putative arms race is completely lacking. The historical dynamics of an arms race

between maternal and paternal alleles can be revealed by assessing how complete PGE is in the hybrid offspring of crosses between closely related species. For example, cryptic sex ratio often reappear in hybrids, free from the constraint imposed by fixed suppressors that have evolved in their original population or species to contain these meiotic drivers (Frank 1991; Hurst & Pomiankowski 1991; Hurst & Werren 2001; Tao *et al.* 2007). Likewise, paternal adaptations against PGE could be unmasked when exposed to divergent maternal backgrounds. The mealybug *Planococcus citri* (Hemiptera: Pseudococcidae) is a particularly-well suited system for such an approach. *P. citri* has emerged in recent years as a model organism for PGE (Brown & Nur 1964; Bongiorni & Prantera 2003; Khosla *et al.* 2006; Ross *et al.* 2010a; Prantera & Bongiorni 2012) and hybridizes readily with other closely related species. A recent study by Kol-Maimon *et al.* (2014a) using hybrid crosses between *P. citri* and *P. ficus* found instances of occasional transmission of the paternal ribosomal ITS2 region through hybrid males, but the presence of hybrid genotypes in their parental *P. citri* population–a result of hybridization in the wild (Kol-Maimon *et al.* 2014b)– and differential amplification in males and females complicates interpretation of their findings. Conclusive evidence requires a larger number of independent genetic markers that allow determining species identity of parental genomes unambiguously: microsatellite loci, now available as a diagnostic tool to distinguish between these two species (Martins *et al.* 2012), are a more suitable tool to confirm whether victory of maternal genomes is complete, or paternal genomes have not yet had their final say.

In this chapter, I aim to empirically test for the existence of an evolutionary arms race between parental genomes in PGE species using a three-generation family study with wild-derived laboratory lines of *P. citri* and *P. ficus* to test two key predictions. First, that paternally-derived chromosomes can escape elimination in males when exposed to a maternal genomic background they have not coevolved with, thus revealing that there is scope for intragenomic conflict between maternal and paternal alleles. On the contrary, if elimination of paternal alleles is found to be complete, such conflict could not be inferred. Second, that escapes occur at a higher rate in hybrid males produced in interspecific crosses between *P. citri* and *P. ficus* than in males produced in intraspecific crosses, consistently, as discussed above, with known examples of reappearance of cryptic drive. This outcome would indicate the existence of independently-evolving anti-

PGE adaptations and maternal suppressors within these two species and, therefore, a historical arms race between parental genomes over elimination of paternal genes. On the other hand, the opposite result—higher leakage in intraspecific crosses—would be consistent with the emergence of drivers and suppressors within current *P. citri* populations. Here, two strategies are employed to detect patrilineal transmission: a panel of polymorphic microsatellite markers (Martins *et al.* 2012) and male response to sex pheromones–a heritable species-specific phenotype (Kol-Maimon *et al.* 2014a; Kol-Maimon *et al.* 2014b). I find that F0 paternal alleles can reappear in F2 genotypes in both hybrid and intraspecific crosses at a similar low frequency, which is consistent with patrilineal inheritance, but no evidence for paternal transmission of pheromone preferences. I therefore conclude that there may be scope for conflict between parental genomes under PGE due to incomplete effectiveness of the mechanism of paternal chromosome exclusion during spermatogenesis, but no clear indication of a recent coevolutionary arms race between parental genomes in these mealybug species.

2.3. Materials and methods

2.3.1. Experimental populations and laboratory rearing

All the experimental crosses in this study were conducted between individuals from three *P. citri* and two *P. ficus* isofemale lines originated from natural populations and reared in the laboratory under a sib-mating regime (Table 2.1). The three *P. citri* lines had undergone at least 15 generations of sib-mating prior to these experiments. Two lines, WYE3-2 and BGOX3, derive from English populations and the third line, CP1-2, originates from Israel. Both *P. ficus* lines (PF1-1 and PF3-1) were derived from Israeli populations and had undergone >8 generations of sib-mating. Mealybug lines were reared on sprouted potatoes placed on tissue paper in sealed containers (boxes or glass/plastic stock bottles) at >50% relative humidity and temperatures of 24-26°C (for *P. citri*) or 26-29°C (for *P. ficus*). To minimize chances of cross contamination, both species were kept in separate rooms. Experimental crosses were kept at 25°C and a 16h-light/8h-dark photoperiod without humidity control.

Table 2.1. Left, mean heterozygosity (observed, H_O, and expected, H_E) and inbreeding coefficient (F_{IS}, Weir and Cockerham 1984) per microsatellite locus of *P. citri* and *P. ficus* lines used in these experiments. For these statistics, only genotypes from polymorphic loci within each line (pol. loci) were considered (see section 2.3.3 for loci information). Estimates were obtained using the online version of Genepop version 4.2 (Raymond & Rousset 1995; Rousset 2008). Right, genome-wide estimates of within-species nucleotide diversity (π) and nucleotide divergence (d_{xy}) and fixation index (F_{ST}) between *P. citri* (WYE3-2 line) and *P. ficus* (PF1-1 line). Estimates were obtained using the blockTools toolkit (github.com/DRL/blocktools). Information about whole-genome sequences used for these estimates is provided in sections 2.3.6 and 3.2.2.

Experimental lines					**Species diversity and divergence**	
Line	Polymorphic loci	H_O	H_E	F_{IS}	$\pi_{P. citri}$	0.0038
WYE3-2	Pci-7, Pci-16, Pci-17	0.159±0.056	0.392±0.086	0.701±0.083	$\pi_{P. ficus}$	0.0045
BGOX-3	Pci-7, Pci-16, Pci-17	0.033±0.009	0.127±0.008	0.770±0.063	d_{xy}	0.0388
CP1-2	Pci-7, Pci-16, Pci-17	0.075±0.012	0.204±0.017	0.648±0.042	F_{ST}	0.8062
PF1-1	Pci-22	0.125	0.525	0.774		
PF3-1	Pci-7, Pci-22	0.313±0.084	0.475±0.022	0.292±0.216		

2.3.2. Experimental crosses

The same experimental cross design was followed in all the experiments in this study and is schematized in Fig. 2.1A. Males and females from different parental F0 lines were isolated and mated to produce F1 cohorts with divergent maternal and paternal haploid genomes. For hybrid crosses, I set 4 biological replicates (i.e. mating pairs) of all possible reciprocal combinations between two *P. citri* lines (WYE3-2 and CP1-2) and the *P. ficus* lines (PF1-1 and PF3-1) and raised the F1 hybrid broods until adulthood. However, I found extremely high levels of hybrid male mortality during early larval stages when crossing *P. ficus* females and *P. citri* males, so that all hybrid males from this genotype (FC hybrids) failed to reach reproductive maturity. This high mortality occurred in all crosses with *P. ficus* mothers, regardless of parental lines; therefore, only allele transmission patterns through hybrid males from *P. citri* mothers (CF hybrids) could be estimated. When possible, I mated 4 CF hybrid males from each hybrid brood to a female from the second line of the maternal species *(P. citri)* to produce F2 offspring (Fig. 2.1B). A simplification of this scheme was used to analyse sex pheromone response

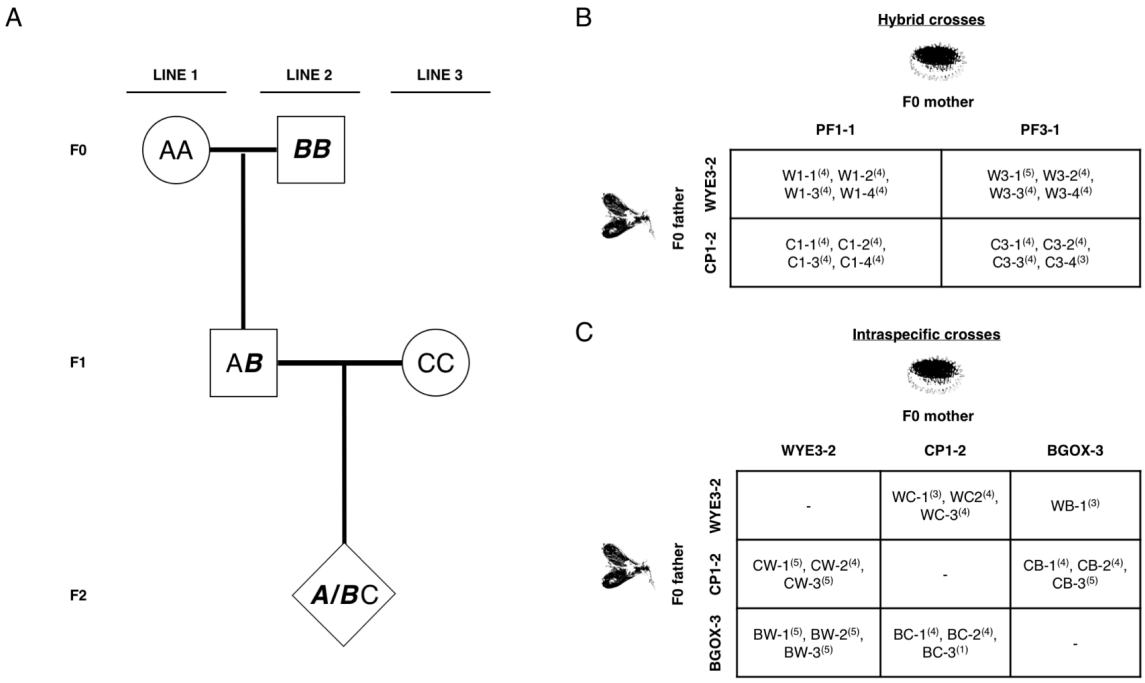


Figure 2.1. (A) Schematic diagram of crossing design. For both hybrid and intraspecific crosses, a female (circle) from an isofemale line with AA genotype at a given locus was mated to a male (square) from a different line (BB) to produce an F1 brood with AB genotype. F1 males from these broods were mated to a female (CC) to produce F2 broods. F1 male transmission ratios were calculated as frequency of maternal allele A in the F2 offspring. Under complete PGE, only AC genotypes are expected in the F2 offspring, so the presence of BC individuals reveals escapes of paternal alleles through F1 males. **(B)** Hybrid crosses and genotypes of F1 broods. 4 biological replicates were produced for each F0 cross. The number of hybrid males from each F1 brood mated to produce F2 families is indicated in brackets. **(C)** Intraspecific *P. citri* crosses and genotypes of F1 broods. 3 biological replicates were produced for each F0 cross except for WB crosses. The number of hybrid males from each F1 brood mated to produce F2s is indicated in brackets.

transmission by hybrid F1 males (see below). For intraspecific crosses, I set 3 biological replicates of all possible reciprocal combinations between the three *P. citri* lines (Fig. 2.1C), raised the F1 broods until adulthood and mated F1 males to a female, preferentially from the parental line that had not been used in the F0 cross.

For all experimental crosses, virgin females were isolated after becoming sexually differentiated (third instar) and kept in separate containers until reproductive maturity (>35-day old). Males were isolated after pupation and kept in clear glass shell vials until emergence of sexually mature adults. Hybrid crosses took place in 6cm-diameter glass Petri dishes with the aid of synthetic pheromones from the paternal species (see below) and occurrence of mating was visually monitored. After mating concluded, the male-female pair was transferred to shell vials containing a single potato

sprout and sealed with cotton wool. For intraspecific *P. citri* crosses, male-female adult pairs were placed directly into shell vials. In both cases, the mating pair was kept in the vial for 3-5 days until egg-laying was observed. Then, males were immediately frozen at -20°C for genotyping and females were transferred to a new rearing container and left to lay eggs for at least 10 days or until death, after which they were removed and frozen at -20° after removal of their bottom half (to avoid genotyping of remaining unlaid eggs).

2.3.3. Microsatellite genotyping

Total genomic DNA from F0, F1 and F2 individuals from experimental crosses was extracted using prepGEM Insect kit (ZyGEM, New Zealand) in 96-well plates following the manufacturer's instructions but reducing the reaction volume by 50%. F0 parents and F1 males were collected for genotyping as described above; F2 individuals were genotyped once they reached the second larval instar. In rare cases where F1 individuals exhibiting genotypes incompatible with parental alleles were observed, accidental contamination was assumed and those crosses were discarded. When possible, females mated to F1 males were also genotyped for intraspecific crosses: otherwise, assignation of parental alleles in the F2 was done under the assumption of complete PGE. In rare cases where individuals exhibiting genotypes incompatible with parental alleles were observed, accidental contamination was assumed and those crosses were discarded.

For genotyping, microsatellite primers for PCR amplification were obtained from Martins *et al.* (2012). A panel of 6 multiplexed loci (Pci-7, Pci-16, Pci-17, Pci-21, Pci-22 and Pci-24) was used in hybrid crosses. For intraspecific *P. citri* crosses, the informative locus panel consisted on the three loci showing intraspecific variability (Pci-7, Pci-16 and Pci-17) and two additional monomorphic loci (Pci-21, Pci-22) to help diagnosing genotyping success for each reaction (Appendix 2, Table S1). PCR amplification of microsatellite loci was performed with Type-it Microsatellite PCR kit (QIAgen, The Netherlands) in a 10 µl reaction volume containing 1 µL of prepGEM reaction product, 5 µL of 2x Master Mix, 0.25 µM of the reverse primer and 0.25 µM of each 5' fluorescently-tagged primer. PCR reactions were performed under the following conditions: initial denaturation at 95°C for 5 min, 32 cycles of denaturation at 95°C for 30s, annealing at 55°C for 90s and extension at 72°C for 30s and a final extension step at 60°C for 30min.

1 µl of PCR product was sent to Edinburgh Genomics for microsatellite genotyping on the ABI 3730 DNA Analyzer system (ThermoFisher Scientific, United Stated of America) with LIZ 500 as size standard. Microsatellite peaks were scored using the Microsatellite Plugin implemented in Geneious 8.1.3 (Biomatters Ltd., New Zealand) and corrected manually.

2.3.4. Allele transmission ratios

For each F2 family and locus, transmission ratios of F1 males were calculated as the proportion of maternally-derived alleles they transmitted to their offspring: i.e. number of occurrences in F2 families of their maternally-inherited alleles divided by total number of F2 genotypes. A transmission ratio of 1 is indicative of complete PGE, while 0.5 denotes Mendelian transmission. For each ratio, an exact binomial test to detect significant deviation from Mendelian expectations was performed in R 3.2.4 (R Development Core Team). To correct for multiple testing, I considered a reduced significance level of $\alpha=0.01$. When possible, transmission ratios for females mated to F1 males were also estimated to confirm Mendelian transmission by calculating the proportion of one of the two alleles (chosen at random) in a heterozygous maternal locus passed on to the F2 offspring.

2.3.5. Sex pheromone response analysis

Interspecific crosses to produce F1 hybrid males were conducted as described above, but using a single line from each species only (WYE3-2 and PF1-1). Due to high mortality of F1 hybrid males with *P. ficus* mothers (FC hybrids), only males from WYE3-2 mothers and PF1-1 fathers (CF hybrids) could be used to produce F2 broods. Ten F0 interspecific crosses were carried out to produce F1 hybrid broods, from which 20-30 males were isolated and mated to a female from the maternal line (WYE3-2) to produce F2s (CF x C), which were raised until adulthood. Intraspecific crosses to produce broods of pure *P. citri* and *P. ficus* males were conducted in an identical way.

Male response trials to both *P. citri* and *P. ficus* sex pheromones were conducted for F0 and F2 males. Synthetic pheromones were provided by Prof. Jocelyn Millar (University of California Riverside) and diluted in 100% ethanol to a concentration of

10ng/µl. The synthetic *P. citri* pheronome (C-phe) used in this experiment is the pure RR enantiomer of (S+)-cis-(1R)-3-isopropenyl-2,2-imethylcyclobutanemethanol acetate, its single component (Bierl-Leonhardt *et al.* 1981). The *P. ficus* pheromone (F-phe) is the racemic component (S)-lavandulyl senecioate (Hinkens *et al.* 2001).

Males were isolated after pupation and kept in shell vials until adulthood. Trials were conducted 24h after adults had emerged from their cocoons in 6cm-diameter glass Petri dish arenas. These arenas contained two 1cm^2 filter paper squares set on opposite sides of the plate, which were randomly impregnated with either 10 ng of pheromone or 1 µl of 100% ethanol (as control). Males were individually placed in the centre of the arena and their responses to both pheromone and control papers were recorded for 15 min. Time of contact with pheromone and control was defined as the number of seconds during which males had any part of their body touching each filter paper. After 15 min, males were taken back to the shell vial for 5 min and then transferred to a second area containing the other pheromone. Time to first contact with pheromone (number of seconds until a male arrived at the pheromone paper for the first time since start of trial) was also recorded. Trials were blind regarding identity and genotype of the males and the two pheromones were tested in a random order.

To analyse total contact times, I corrected for time spent on the control paper during trials by subtracting the number of seconds males were in contact with the control from the number of seconds in contact with the pheromone in each trial. Negative values of this corrected measurement (i.e. when a male spent more time on the control paper than the pheromone) were given a value of 0 as I considered that these males did not show a true pheromone response. I fitted a series of mixed models using the 'lme4' R package (Bates *et al.* 2015) to test whether patterns of pheromone response differ between the three groups of males included in this study (*P. citri*, *P. ficus* and CF x C F2 offspring). First, I fitted a binomial GLMM to test for differences in the frequency of responding males to both pheromones across genotypes. Then, I fitted two linear mixed models to further explore two additional aspects of behaviour of responding males: intensity of attraction (total time in contact with pheromones) and speed of response (time to first contact). In all three models, I included pheromone, genotype and their interaction as fixed effects. I also included order of exposure to both pheromones as an additional fixed effect, and male

ID as a random effect. I used likelihood ratio test to assess significance of fixed effects and Tukey *post hoc* comparisons to test for differences between pairs of genotypes using the 'multcomp' R package (Hothorn *et al.* 2008)

2.3.6. Species confirmation and primer mapping

In order to confirm species identity of the WYE3-2 and PF1-1 lines used in the sex pheromone response experiment, I retrieved the 28S–D2, ITS2, COI–region 2 and COI–LCO sequences from the genome assemblies generated for both lines by our research group (PCITRI.V1 and PFICUS.V0, publicly available in http://mealybug.org). To obtain these sequences, I blasted the *P. citri* sequences for those regions obtained by Malausa *et al.* (2011) against both assemblies using the BLAST tool in mealybug.org with default settings. The best matches from each species were compared to GenBank sequences using the NCBI BLAST tool (http://blast.ncbi.nlm.nih.gov). To reveal the extent of genome coverage of our microsatellite panel, I mapped all loci against both genome assemblies. All forward and reverse primer sequences were blasted against PCITRI.V1 and PFICUS.V0 using the BLAST tool in mealybug.org with default settings.

2.4. Results

2.4.1. Microsatellite panel optimization

I initially tested eight markers (Pci-6, Pci-7, Pci-14, Pci-16, Pci-17, Pci-21, Pci-22 and Pci-24) for amplification in at least 3 individuals from all *P. citri* and *P. ficus* lines. All markers successfully amplified in the three *P. citri* lines, while Pci-6 and Pci-14 failed to amplify in *P. ficus.* Since these two loci were found to be monomorphic in our *P. citri* lines, they were discarded for further genotyping.

A list of the alleles amplified in both species is provided in Table S1 of Appendix 2. BLAST searches revealed that all markers are located within different scaffolds in both genome assemblies. Pci-16, Pci-17, Pci-21, Pci-22 and Pci-24 were found to be optimal diagnostic markers for hybrid crosses due to the presence of species-specific alleles

(Supplementary Fig. S1, Appendix 2). Pci-7 was also included in the genotyping panel for hybrid crosses due to its high allelic richness, even when alleles were shared between both species. For intraspecific crosses, only Pci-7, Pci-16 and Pci-17 were found to be polymorphic within *P. citri.*

2.4.2. Allele transmission in hybrid crosses

In order to test allele transmission in hybrid males, I initially aimed to conduct all reciprocal crosses between both *P. citri* and *P. ficus* lines. However, I found extreme sex-specific mortality in crosses with *P. ficus* mothers and *P. citri* fathers: hybrid males from these crosses (FC hybrids) consistently failed to reach adulthood, regardless of parental lines or raising conditions. I set at least 4 replicates of each FC cross from all possible combinations (PF1-1 and PF3-1 mothers x WYE3-2 and CP1-2 fathers) and only obtained three adult males, none of which managed to successfully inseminate a female to produce F2 broods. Therefore, all hybrid males that survived to adulthood and fathered F2 broods in this study derive from CF crosses (WYE3-2 and CP1-2 mothers x PF1-1 and PF3-1 fathers).

All F0 *P. citri* mothers and *P. ficus* fathers were genotyped to confirm the presence of alleles specific to both species at the diagnostic loci (Pci-16, Pci-17, Pci-21, Pci-22 and Pci-24) and to determine their Pci-7 genotypes. All F1 broods were genotyped after mating to confirm expected genotypes in case of parental homozygosity and to determine their genotypes at those loci they were heterozygous for in one or both F0 parents. In all cases of parental heterozygosity, F1 genotypes adjusted to expected patterns of PGE transmission: heterozygous F0 males transmitted one allele only (that of maternal origin) to all genotyped F1 males and heterozygous F0 females transmitted both. Rarely, genotyping of F1 hybrid males unexpectedly revealed *P. citri* alleles only, which was interpreted as instances of accidental contamination of F1 hybrid bottles with males from the maternal species and led to discarding of whole affected broods.

From each F1 brood, 4 hybrid males were mated to father 4 F2 families each (with two exceptions: 5 males for W3-1 and 3 males from C3-4), yielding 64 F2 families in total. I found that 63/64 F1 hybrid fathers showed complete PGE (Fig. 2.2A): they only transmitted maternally-derived alleles to the F2 at all loci and no grandpaternal (i.e. *P.*

ficus) alleles were found in the genepools of the F2 broods they fathered. The only exception was W1-4_1 (Appendix 2, Table S2). This male transmitted his paternally-inherited allele at two loci to the same F2 individual (Pci-17 and Pci-22). The other loci in this individual showed maternal alleles only, as expected under PGE. The remaining 11 genotyped individuals fathered by W1-4_1 received his maternal alleles only at these loci.

In total, 356 transmission ratios across all males and loci were estimated (at least 5 for each male at the 5 inter-species diagnostic loci and an additional one at Pci-7 for 36 males which were also heterozygous at that locus). 354 of these ratios, all except for these two exceptions mentioned before, had a value of 1, indicating complete PGE. Assuming equal probability of transmission of paternal alleles across all loci, I obtained an estimation of frequency of paternal escapes of 0.0007-0.0201 (95% CI). Even though I did not genotype females mated to F1 males, it was also possible to estimate maternal transmission ratios when a *P. citri* allele different from the one transmitted by the F1 father was observed in F2 families. These cases were indicative of maternal heterozygosity and allowed to determine whether F1 females transmitted alleles in a Mendelian way, as expected in a PGE system. 27 transmission ratios for F1 females were estimated, none of which deviated significantly from Mendelian expectations ($p>0.01$).

2.4.3. Allele transmission in intraspecific crosses

For intraspecific crosses, three biological replicates were set for each possible cross. However, only one replicate of the cross between WYE3-2 mothers and BGOX-3 father could be successfully raised into adulthood. For each F1 brood, between 3 and 5 males were mated to produce F2 broods. In total, I obtained transmission patterns for 65 F1 males at least at one informative locus (Fig. 2.2B). Of these, 3 F1 males showed allele transmission patterns consistent with incomplete paternal genome elimination (Appendix 2, Table S2). BW_2_3 and BC1_3 transmitted a paternal allele once, both at the Pci-7 locus, and CB_3_1 passed on paternal alleles to the same F2 individual at the Pci-7 and Pci-17 loci. I validated all these exceptions by re-genotyping the F2 individuals showing escaped alleles. However, I could not genotype the female that was mated to BC1_3, which opens the possibility that the seemingly paternal allele that would have

been transmitted by this male to one of their offspring actually derives from the F1 mother. Since the hypothesis of Mendelian inheritance for the putatively escaped allele at that locus cannot be rejected at the reduced significance level used in this study (11 F2 individuals, p=0.012), this escape cannot be unambiguously confirmed.

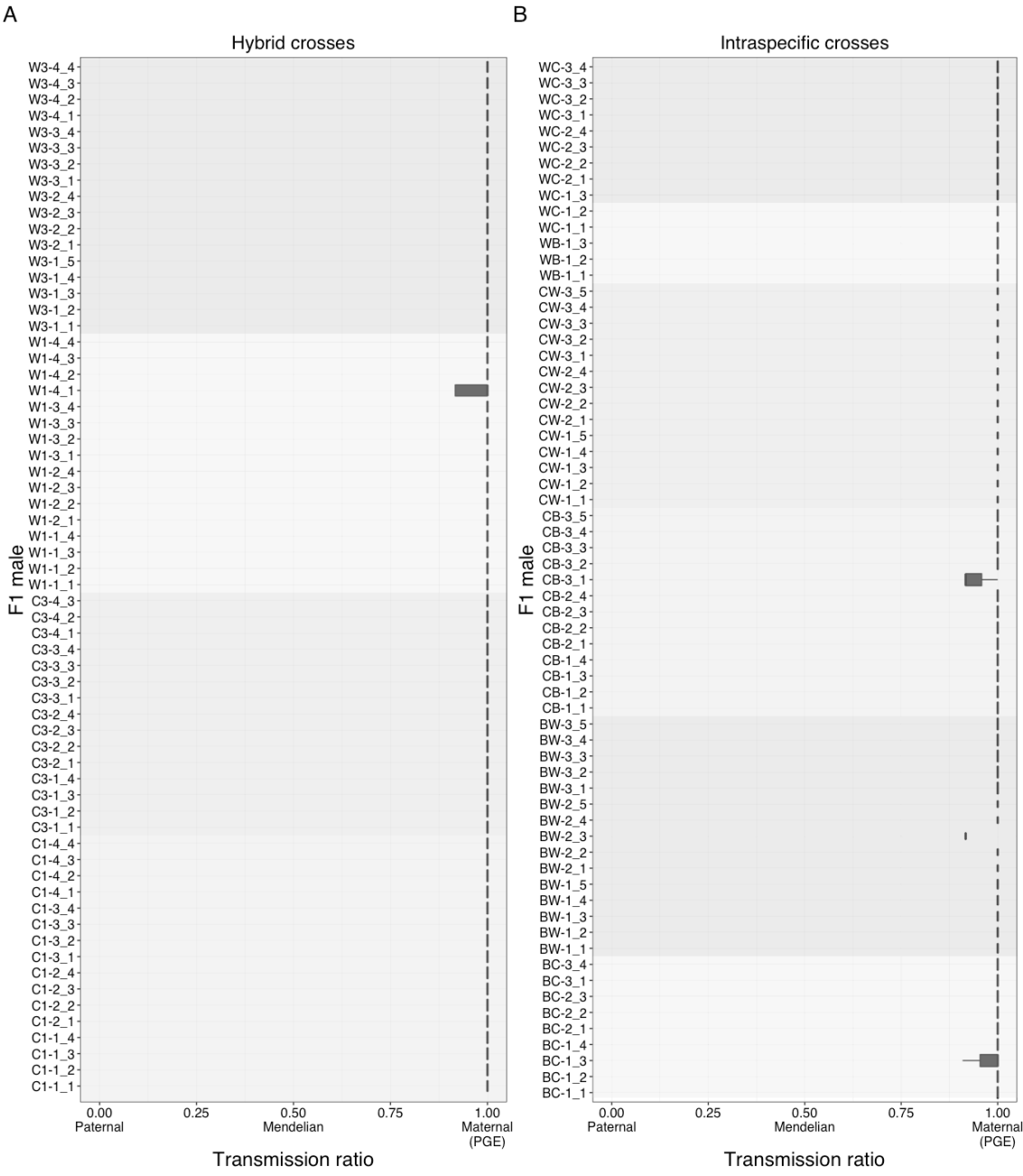


Figure 2.2. Paternal allele transmission ratios for F1 males in hybrid **(A)** and intraspecific *P. citri* crosses **(B)**. F1 males are labelled as follows: the first two characters denote F0 maternal and paternal genotypes, followed by a number corresponding to the F0 cross and a second number indicating the identity of the male: e.g. C1-1_1 refers to the first F1 male deriving from the first CP3-2 x PF1-1 F0 cross.

Overall, 144/148 transmission ratios across all males and loci were consistent with complete PGE. As in hybrid crosses, I also estimated 21 transmission ratios for F1 mothers, none of which significantly departed from Mendelian expectations. The frequency of paternal escapes in intraspecific crosses, again assuming equal rates across all three loci, was estimated to be 0.0074-0.0678 (95% CI). This value was not significantly different to the estimation obtained for hybrid crosses (Pearson's χ^2 test with Monte Carlo simulation, p=0.065). Combining data from hybrid and intraspecific males, I obtained a common estimation of paternal escapes frequency of 0.0044-0.0257 (95% CI). In total, 4 of 1,548 genotyped F2 individuals between hybrid and intraspecific crosses received an escaped allele.

2.4.4. Response to sex pheromones

To test whether F1 hybrid males can transmit paternal pheromone preferences to the offspring, I tested and compared response patterns to both C-phe (*P. citri)* and F-phe (*P. ficus)* pheromones between groups of pure males from both species and F2 offspring of CF fathers and *P. citri* mothers. Under complete PGE, CF males should always transmit *P. citri* (i.e. maternal) pheromone preferences to their sons and therefore males from F2 broods should exhibit identical pheromone responses to pure *P. citri* males.

As expected, most pure species males showed a response toward their conspecific pheromone (86.3% of *P. citri* males responded to C-phe and 80.1% of *P. ficus* males responded to F-phe). However, I also found strong cross-attraction to the pheromone from the other species (47.1% of *P. citri* males were attracted to F-phe and 53.8% of *P. ficus* males to C-phe). The F2 offspring of CF fathers crossed to *P. citri* mothers showed similar responses to *P. citri* males: 82.7% of them were attracted to C-phe, while 56.5% responded to F-phe. (Fig. 2.3A). Attraction to both pheromones was showed by 41.2% of *P. citri* males and 53.8% of *P. ficus* males, while 7.8% and 19.2% failed to respond to either pheromone respectively. The frequencies of F2 males that showed response to both pheromones (49.3%) and lack of response to either (10.1%) were similar to *P. citri* males (Fig. 2.3B).

A series of mixed models were used to test for differences in sex pheromone response patterns across the three different genotypes (Appendix 2, Table S3). First, I fitted a binomial mixed model to detect significant differences in proportion of males from each genotype that responded to C-phe and F-phe. I found a significant effect of the interaction between genotype and pheromone ($LR_{8,6}$=20.59, p<0.001). The order in which males were exposed to the pheromones did not have an effect on response ($LR_{8,7}$=0, p=0.997). Male identity, fitted as a random effect, explained 19% of the variance in response (σ^2_{ID}=0.186). *Post hoc* comparisons revealed significant differences in pheromone response between intraspecific males from both species: *P. citri* males show stronger response to C-phe than *P. ficus* males (Z=-2.856, p=0.047) but, conversely, *P. ficus* males are not more strongly attracted to their own pheromone than *P. citri* males (Z=2.669, p=0.079). Comparisons between how these two genotypes responded to both pheromones revealed a similar pattern: *P. citri* males are more attracted to C-phe than to F-phe (Z=-3.933, p=0.001), but there is no significant difference in attraction to either pheromone in *P. ficus* males (Z=2.120, p=0.270). I found no significant difference in response to either pheromone between *P. citri* males and F2 males (C-phe: Z=-0.374, p=0.999; F-phe: Z=1.269, p=0.796).

Second, a linear mixed model was fitted to test whether there was any difference across genotypes in intensity of attraction, represented by the total time spent by responding males in contact with the sex pheromones (Fig. 2.3C). Again, I found a significant interaction between genotype and pheromone ($LR_{9,7}$=13.443, p=0.012) and no effect of order of exposure ($LR_{9,8}$=0.193, p=0.660). The proportion of the variance explained by male identity was estimated to be 11%. *Post hoc* comparisons revealed that both *P. citri* and F2 responsive males spent more time in contact with C-phe than F-phe (Z=2.120, p<0.001 and Z=-5.120, p<0.001), but *P. ficus* males did not show a significant difference in contact time with either pheromone (Z=0.766, p=0.972). No significant differences in contact time with either C-phe or F-phe were found between *P. citri* and F2 males.

Finally, a second linear mixed model was used to test for differences in speed of response across genotypes (Fig. 2.3D). In contrast with previous models, I did not find any significant difference in time to first contact with the pheromones across the three genotypes: there was no significant effect of an interaction between genotype and

pheromone ($LR_{9,7}$=3.885, p=0.143). Again, order of exposure to pheromones had no significant effect either ($LR_{9,8}$=0.901, p=0.343). Male identity explained 19% of the variance. Together, these models revealed no difference between *P. citri* and F2 males, indicating that CF males were not able to transmit paternal pheromone preferences to their offspring.

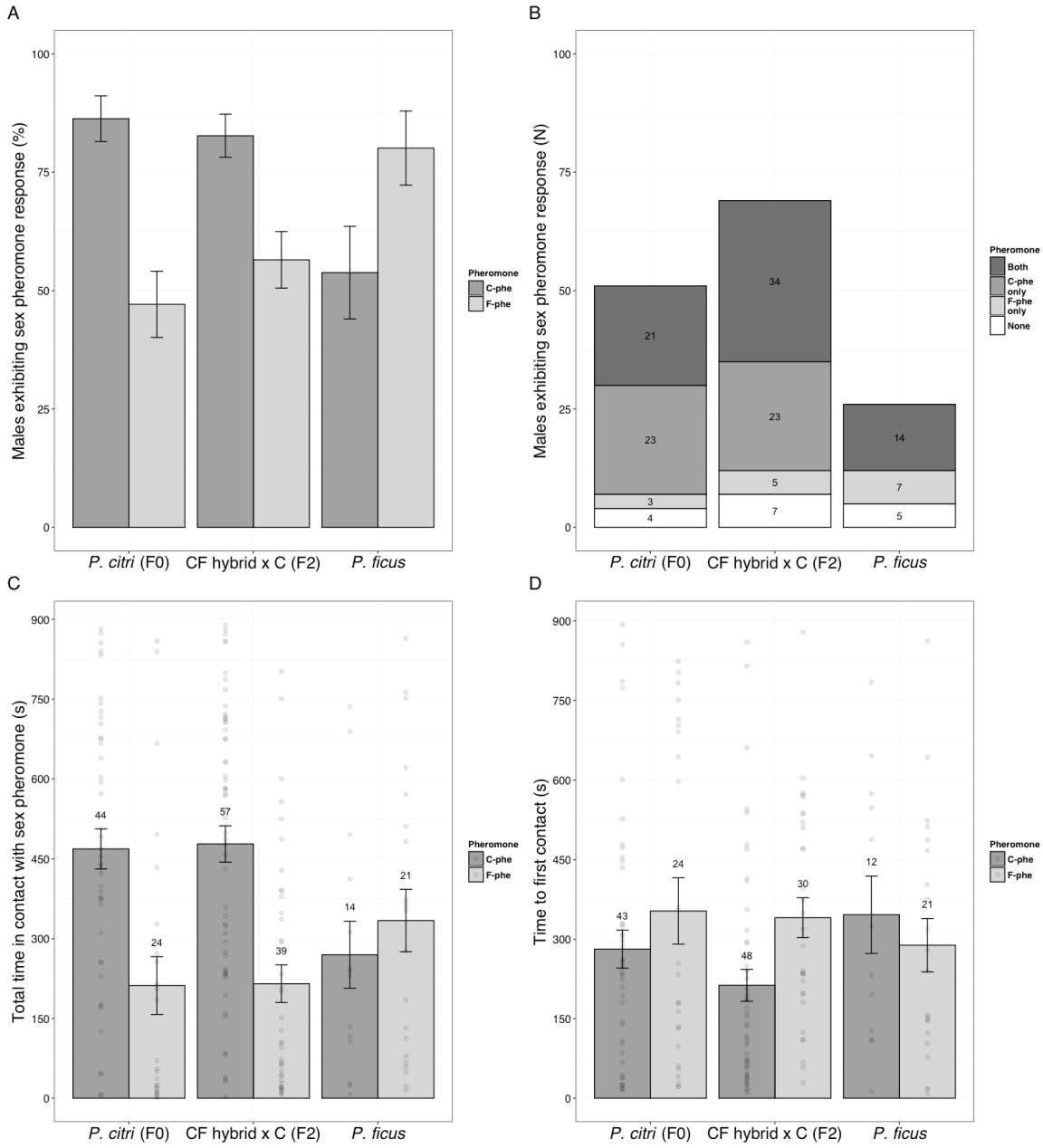


Figure 2.3. Male response to sex pheromones. **(A)** Percentage of males of each genotype responding to both C-phe and F-phe. **(B)** Number of males exhibiting attraction to both pheromones, either or none. **(C)** Number of seconds spent by responding males in contact to both pheromones. **(D)** Time to first contact of responding males. Error bars represent standard errors (binomial standard error in panel A). Number of males exhibiting pheromone response from each genotype is shown above error bars in C and D.

Second, a linear mixed model was fitted to test whether there was any difference across genotypes in intensity of attraction, represented by the total time spent by responding males in contact with the sex pheromones (Fig. 2.3C). Again, I found a significant interaction between genotype and pheromone ($LR_{9,7}$=13.443, p=0.012) and no effect of order of exposure ($LR_{9,8}$=0.193, p=0.660). The proportion of the variance explained by male identity was estimated to be 11%. *Post hoc* comparisons revealed that both *P. citri* and F2 responsive males spent more time in contact with C-phe than F-phe (Z=2.120, p<0.001 and Z=-5.120, p<0.001), but *P. ficus* males did not show a significant difference in contact time with either pheromone (Z=0.766, p=0.972). No significant differences in contact time with either C-phe or F-phe were found between *P. citri* and F2 males.

Finally, a second linear mixed model was used to test for differences in speed of response across genotypes (Fig. 2.3D). In contrast with previous models, I did not find any significant difference in time to first contact with the pheromones across the three genotypes: there was no significant effect of an interaction between genotype and pheromone ($LR_{9,7}$=3.885, p=0.143). Again, order of exposure to pheromones had no significant effect either ($LR_{9,8}$=0.901, p=0.343). Male identity explained 19% of the variance. Together, these models revealed no difference between *P. citri* and F2 males, indicating that CF males were not able to transmit paternal pheromone preferences to their offspring.

2.5. Discussion

Paternal genome elimination is a genetic system characterised by whole-genome meiotic drive of maternally-inherited chromosomes in males at the expense of paternally-inherited homologs. Because of this extreme deviation from fair Mendelian inheritance, PGE is expected to generate intragenomic conflict between maternal and paternal haploid genomes within males. The evolutionary success of PGE, which has independently emerged several times in Arthropoda and is estimated to be present in over 10,000 species (Burt & Trivers 2006; Gardner & Ross 2014; de la Filia *et al.* 2015), suggests that this conflict has been irrevocably resolved in favour of maternal genomes.

Yet this notion seems difficult to reconcile with the dramatic differences in timing of elimination and degree of expression of paternal chromosomes observed not only across PGE origins, but also between closely related species (Normark 2003; Ross *et al.* 2010a). Verbal models have predicted a coevolutionary arms race between parental genomes under PGE, triggered by strong selection on the paternal genome to escape elimination and subsequent maternal counteradaptations (Herrick & Seger 1999; Burt & Trivers 2006; Ross *et al.* 2010a). In this study, I aimed to determine whether there is evidence for such an arms race by confronting independently-evolving maternal and paternal genomes within males produced in hybrid and intraspecific crosses. I tracked the inheritance of both a genotypic (microsatellite markers) and a phenotypic trait (sex pheromone response) to determine if these males exhibited incomplete PGE consistent with a mismatch between parental genomes. The results of these experiments reveal that elimination of paternally-derived chromosomes is not completely effective, suggesting scope for intragenomic conflict, but do not offer enough evidence to infer the existence of an arms race between parental genomes.

2.5.1. Detectable instances of transmission of paternal chromosomes through males but no evidence of a coevolutionary arms race between parental genomes

Escapes of paternal alleles revealing incomplete PGE could be detected in both hybrid and intraspecific males. Although these escapes were relatively few, they are far from negligible considering the limitations of a classical microsatellite approach with a limited number of diagnostic markers. I genotyped up to 12 F2 offspring per cross, less than 5% of the average number of eggs laid by *P. citri* females in experimental conditions (300-500 eggs) (Myers 1932; Ross et al 2010b). Also, I could only use three informative markers for intraspecific crosses, which falls short of covering the haploid complement of these species (n=5) (Hughes-Schrader 1948). Even so, I could detect escapes at a frequency of 0.4-2.5%, which is substantial at the population scale. If anything, this study is likely to underestimate escapes due to partial genome coverage and low offspring number that can be feasibly genotyped with such a design. As for the existence of a coevolutionary arms race between parental genomes in mealybugs, these results are inconclusive. I did not find a higher frequency of escapes in hybrids than in intraspecific males, which would be indirect evidence for historical coevolution of paternal and

maternal genomes. Also, for such an arms race to occur, the ability of paternal alleles to escape elimination must be heritable. The experimental design does not allow determining whether escapes are accidental or if there is a heritable component to incomplete PGE, which would require a larger multi-generational crossing design.

The observed leakages of paternal alleles cannot be explained by recombination, since meiosis is achiasmatic in mealybug males (Bongiorni *et al.* 2004), so it must be attributed to transmission of entire paternal chromosomes. Interestingly, even when half of the males exhibiting incomplete PGE transmitted a paternal allele at more than one locus at once, only maternal copies were transmitted at other informative loci. This indicates that paternal escapes can affect one or more chromosomes at the same time (our genome assemblies are not complete enough to assign markers to chromosomes, so loci affected by paternal transmission simultaneously could be situated on the same chromosome), but not all. I did not find a clear pattern across loci suggesting differences in likelihood to escape elimination either. Most paternal escapes were found at Pci-7, but our detection power was highest for this locus. Also, these results cannot unambiguously reveal whether escaped paternal chromosomes are transmitted in addition to their maternal homologs or by replacing them. In all cases of paternal leakage, the allele transmitted by to the F2 offspring by their mother and the maternally-inherited allele in the F1 father were the same (Appendix 2, Table S2), so triploid microsatellite peaks resulting from transmission of both parental homologs could not be found.

However, due to the low number of F2 individuals found to receive a paternal allele in this study, these results must be interpreted with caution. The rate of F2 genotypes with an escaped allele found in this study is in the order of 10^{-3}, which is in the range of typical error rates in microsatellite studies (Pompanon *et al.* 2005; Hoffman & Amos 2005; Guichoux *et al.* 2011) and higher frequency SSR mutation rates (Ellegren 2000). Common causes of reproducible error, such as null alleles or allelic dropout (Dakin & Avise 2004), can be confidently excluded, as all escapes resulted in heterozygous genotypes. Of particular concern in our methodology could be cross-contamination between samples, since DNA extractions were performed in 96-well plates instead of individual tubes. For hybrid crosses, however, the reappearance of diagnostic *P. ficus* alleles in a single individual cannot be explained by contamination, since all other individuals in the same plate exclusively carried *P. citri* alleles transmitted

by their hybrid fathers and their *P. citri* mothers; also, two simultaneous mutations to *P. ficus* alleles affecting the maternal *P. citri* genome in a F1 hybrid would be extremely unlikely. However, independent assessment of paternal escapes through more robust SNP-based parentage methods, which would offer a much higher number of traceable markers (Elshire *et al.* 2011; Kaiser *et al.* 2017) would offer a superior evaluation of paternal chromosome leakages and provide information on their distribution in the genome, which cannot be inferred with this microsatellite panel. In combination with a deeper pedigree, this approach would facilitate to explore heritability of paternal escapes in mealybugs and increase our power to examine the coevolutionary dynamics between maternally- and paternally-inherited genomes.

2.5.2. A likely cause of transmission of paternal chromosomes: sporadic failure of meiotic parent-of-origin discrimination mechanisms

A complete understanding of PGE at the molecular level is still lacking, although available data provides some clues on how paternal chromosome leakages might occur. Mealybug spermatogenesis has been studied extensively and the sequence of events and timing of paternal genome elimination are well described (Hughes-Schrader 1948; Bongiorni *et al.* 2004; Bongiorni *et al.* 2009). In mealybugs, meiosis follows an inverted sequence (Chandra 1962; Viera *et al.* 2008) and segregation of parental homologs is delayed until anaphase II, which involves a monopolar spindle that only interacts with the euchromatic maternal set. Only the spermatids carrying maternal chromosomes progress to complete sperm maturation, while spermatid nuclei containing the paternal set, which lags behind in anaphase II, degenerate (Bongiorni *et al.* 2004). How can paternal chromosomes escape this fate? Several lines of evidence suggest that escapees must undergo a reversal of heterochromatinization to allow attachment of the monopolar spindle. In the mealybug *Pseudococcus viburni*, paternally-inherited material that loses its condensed state during meiosis (either supernumerary B chromosomes or irradiation damaged autosomes) segregate into active sperm with the maternal complement (Nur 1962a; Brown & Nur 1964). Moreover, due to the holocentric nature of mealybug chromosomes (i.e. they lack a localised centromere) (Schrader 1935; Wrensch *et al.* 1994), partial lack of heterochromatinization can be sufficient for spindle attachment. Translocated chromosomes with both euchromatinized maternal and

condensed paternal segments have been shown to migrate preferentially with the maternal set (Nur 1970). If reversal of heterochromatinization is necessary for paternal replacement of maternal chromosomes, either mutations or sporadic failures (or manipulation by the paternal genome) of the epigenetic machinery that encodes parent-of-origin chromosome information in mealybugs during spermatogenesis, such as DNA methylation levels (Bongiorni *et al.* 1999) or histone modifications (Khosla *et al.* 2006; Prantera & Bongiorni 2012) (Khosla *et al.*, 2006; Prantera & Bongiorni, 2012)—which undergo extensive reorganization during meiosis (Bongiorni *et al.* 2009)—, could be responsible for paternal leakages.

2.5.3. Transmission of sex pheromone preferences through CF hybrids follows PGE expectations

I found no evidence of transmission of paternal sex pheromone preferences through males. An important difference between this study and previous work on pheromone response that complicated the predicted outcome of this experiment is cross-attraction to C-phe shown by half of the tested *P. ficus* males, which does not occur in wild populations (Kol-Maimon *et al.* 2014b) and had only been reported before as a rare event in laboratory conditions (Kol-Maimon *et al.* 2010). The reasons for this cross-attraction are unclear and cannot be attributed to contamination of *P. ficus* experimental cultures with *P. citri* males, as I routinely genotyped *P. ficus* individuals used in trials with our diagnostic microsatellite panel and, additionally, confirmed the species identify of the PF1-1 line using common barcoding regions (Appendix 2, Table S4). Nevertheless, statistical analyses did not detect differences in different components of pheromone response between male offspring of CF hybrids and the grandmaternal species, *P. citri*. The genetic architecture of sex pheromone response remains unexplored in mealybugs, yet in other insect species such as moths and *Drosophila* the specificity of male pheromone response has been shown to be controlled by single genes or several tightly linked loci (Roelofs *et al.* 1987; Löfstedt 1993; Kurtovic *et al.* 2007; Gould *et al.* 2010). If this is also the case in mealybugs, the observed leakage of paternal chromosomes could be sufficient for transmission of the genetic toolkit involved in pheromone response to F2 males, but expression of preference would be dependent on overriding silencing of paternal chromosomes.

Patrilineal inheritance of sex pheromone preferences was previously reported by (Kol-Maimon *et al.* 2014a) in an analogous experimental setup using FC males. I was unable to raise viable FC males and could only test sex pheromone responses in F2 broods produced in the reciprocal cross. Since their study did not explore paternal transmission through CF hybrids, both studies may be complementary and suggest a parental species effect in PGE failure, with *P. citri* alleles being more prone to be expressed in an *P. ficus* maternal background than vice versa. This hypothesis is further supported by the strong differences in mortality of reciprocal hybrid males in this and, to a lesser magnitude, previous studies (Rotundo & Tremblay 1982; Tranfaglia & Tremblay 1982). Early condensation of paternal chromosomes during male development should prevent the expression of paternal alleles, as shown by inheritance studies of phenotypic markers in *P. citri*, which are expressed in males when maternally-inherited only and regardless of dominance (Brown & Nur 1964; Brown & Wiegmann 1969). Since maternal genomes are responsible for maintaining paternal chromosomes silencing in mealybugs (Brown & Nur 1964; Ross *et al.* 2010a), a likely explanation for the reproducible failure of FC matings to produce viable sons would be maternal *P. ficus* backgrounds failing to properly regulate silencing of paternal *P. citri* genomes, leading to expression of harmful Dobzhansky-Muller incompatibilities between parental genomes (Orr 1996; Johnson 2010).

2.5.4. Conclusion and future directions

In this chapter, I have obtained empirical evidence for sporadic transmission of paternal chromosomes through mealybug males, but not for sex pheromone preferences through CF hybrids or antagonistic coevolution between parental alleles. These results confirm and complement previous reports of paternal transmission of genetic markers and pheromone responses through FC male hybrids (Kol-Maimon *et al.* 2014a). Taken together, these two studies reveal differences in the ability of reciprocal hybrid mealybug males to transmit and express paternal preferences, which most likely depend on asymmetric interactions between the genomes of these species in a hybrid background. The estimation of leakage rate obtained in this study will be a useful starting point for future broader studies aiming to explore these effects and readdress the existence of an

ongoing arms race between parental alleles in current populations of *P. citri* or its historical occurrence within the genus.

Chapter 3

Broken silence: Parent-of-origin-specific transcriptome analysis of hybrid mealybug males

3.1. Summary

Genetic conflict has been brought up to explain the evolution of non-Mendelian genetic systems in which parents do not contribute equally to the genetic makeup of their offspring. One of the most extraordinary examples of such systems is paternal genome elimination (PGE), a form of haplodiploidy which has independently evolved several times in arthropods. PGE males are diploid but systematically transmit maternally-inherited chromosomes only, while the paternal homologues are excluded from sperm. Due to these asymmetric inheritance patterns, theoretical work suggests an evolutionary arms race between paternal and maternal genomes over transmission to following generations. Consequently, in some PGE species such as the citrus mealybug *Planococcoccus citri*, paternal chromosomes are heterochromatinised early in development and are thought to remain inactive so that possible paternal genome responses to resist its elimination are prevented. Here, I present a parent-of-origin allele-specific transcriptome analysis in hybrid mealybugs from crosses between *P. citri* and the closely related *P. ficus*. I show that expression is globally biased towards the maternal genome but detect activity of paternal chromosomes in both somatic and reproductive tissues. Up to 70% of genes with somatic expression are to some degree paternally-expressed. Our results provide a first insight into patterns of gene expression under PGE and offer a solid ground to further explore if activity of paternal alleles can induce an evolutionary response against whole-genome drive of maternal genomes.

3.2. Introduction

The laws of inheritance tend to be remarkably fair and symmetric in sexual reproduction (Wright 1931). In most sexually reproducing animals, mothers and fathers contribute equally to the genetic makeup of their offspring. Each parent transmits a haploid chromosomal set to both sons and daughters, which are in turn equally likely to express and transmit alleles independently of the parent they inherited them from (Normark 2006). Although these symmetries are ancestral and present in most sexually-reproducing eukaryotes, there are exceptions. A small fraction of the genome is

frequently transmitted in a parent-of-origin dependent manner: for example, genes located in Y chromosomes, which are only transmitted from fathers to sons (Bachtrog *et al.* 2014), or mitochondrial genomes, which are maternally-transmitted (Birky 1995). Moreover, in some species, non-Mendelian segregation can even affect entire haploid genomes: this is the case for certain genetic systems such as parthenogenesis, in which females produce diploid daughters without paternal contribution (Normark 2003), or arrhenotoky (i.e. true haplodiploidy), where haploid males develop from unfertilised eggs and only inherit their mothers' genes (Normark 2006; Normark & Ross 2014). Parent-of-origin-dependent gene expression is also relatively common. For example, imprinted nuclear genes—only expressed when paternally- or maternally-inherited—in mammals and flowering plants (Reik & Walter 2001; Ferguson-Smith 2011), or paternal X chromosome inactivation in marsupials (Deakin *et al.* 2009).

The evolutionary causes of these asymmetries remain poorly understood. A commonly invoked hypothesis is intragenomic conflict (Hurst 1992; Burt & Trivers 2006; Werren 2011; Gardner & Úbeda 2017): under this theory, asymmetries in gene inheritance and expression are the result of the clash between different entities with opposing evolutionary agendas that coexist within genomes (Hurst 1992; Haig 2000; Burt & Trivers 2006; Ross *et al.* 2010a; Werren 2011; Normark & Ross 2014; Gardner & Úbeda 2017). Most empirical work testing these theories has focused on mammals and other model organisms, but the most extreme cases of extreme violations of fair transmission and biparental expression, such as asymmetric genetic systems, have barely been explored. These systems offer a valuable opportunity to understand the role of conflict in breaking ancestral symmetries and the evolutionary consequences of such violations. Here, I focus on paternal genome elimination (PGE), a unique genetic system that combines non-Mendelian inheritance and parent-of-origin-specific asymmetries in gene expression creating strong scope for intragenomic conflict between parental alleles.

PGE is a form of haplodiploidy with seven independent origins across arthropods and is estimated to be present in more than 10,000 species (de la Filia *et al.* 2015). Under PGE, both sexes are diploid—in contrast to arrhenotoky—but males only transmit maternally-inherited chromosomes to their offspring, while the paternally-inherited set fails to be incorporated into sperm (Normark 2003; Burt & Trivers 2006; Gardner & Ross

2014; Blackmon *et al.* 2015). Therefore, gene transmission through males is not random but dependent on whether they are maternally-derived. Moreover, in most PGE taxa, paternal chromosomes are destroyed (e.g. in mites or armored scale insects) or transcriptionally silenced (e.g. in mealybugs or the coffee borer beetle) early in development, so that males express maternally-inherited alleles only (Burt & Trivers 2006; Gardner & Ross 2014; de la Filia *et al.* 2015). Since these asymmetric inheritance patterns confer a transmission advantage to mothers through their sons, the evolution of PGE has been frequently framed as the outcome of intragenomic conflict between parental alleles in males (Bull 1979; Haig 1993; Herrick & Seger 1999; Ross *et al.* 2010a; Normark & Ross 2014).

Under this view, PGE repeatedly emerges when maternal alleles become able to manipulate spermatogenesis in males to enhance their own transmission, and the subsequent evolution of suppression of paternal allele expression (either by destroying or silencing paternal chromosomes) would be favoured due to conferring further maternal advantage, in spite of associated costs to male fitness such as exposure of deleterious recessive alleles (Brown 1964, Gardner & Ross 2014). This maternal benefit from male haploidisation is two-fold: first, it would help reinforcing drive of maternal alleles during spermatogenesis by preventing paternal anti-PGE adaptations from restoring fair Mendelian transmission; second, it could alleviate possible deleterious effects of intragenomic conflict within males (see for example Crespi & Summers 2005; Úbeda & Wilkins 2008; Gardner & Úbeda 2017).

Indirect evidence supports this scenario: for example, in the mealybug *Pseudococcus viburni* (Hemiptera: Pseudococcidae), where paternal chromosomes are heterochromatinized, the presence of the maternal genome is a necessary condition for the maintenance of their condensed state (Nur 1962b; Brown & Nur 1964). However, while embryonic elimination of paternal chromosomes leaves no question as to whether expression of paternal alleles is fully prevented, embryonic silencing is a more complex situation that remains poorly understood. On one hand, in contrast to embryonic elimination of paternal chromosomes, silencing is more susceptible to manifestation of intragenomic conflicts, as putative anti-PGE paternal strategies which managed to interfere with silencing could ultimately threaten whole-genome drive of maternal alleles—either by themselves or by unleashing prior paternal responses to PGE acting

on spermatogenesis. On the other hand, haplodisation by silencing could be expected to carry a reduced cost on male fitness, as condensed paternal chromosomes could be recruited to contribute transcriptional activity, and therefore be of interest to maternal genomes. To date, very little is known about how completely paternal chromosomes are silenced in the male soma in PGE species with paternal chromosome heterochromatinization—in soma, where functional male haploidy could compromise viability of males, and in testis, where paternal genes are ultimately lost and putative paternal anti-PGE adaptations and maternal counterresponses are thus most likely to operate.

Mealybugs are a prime model organism to explore transcriptional suppression of paternal chromosomes and its evolutionary role in parental genome conflict under PGE, as they are arguably the group in which most of the empirical work on cytological and epigenetic manifestations of this system has been focused (Brown & Nelson-Rees 1961; Brown & Nur 1964; Bongiorni *et al.* 1999; Khosla *et al.* 2006; Bongiorni *et al.* 2009; Ross *et al.* 2010a; Prantera & Bongiorni 2012). The genetic inactivity and paternal origin of heterochromatic chromosomes was first shown with radiation experiments in the mealybug *Planoccocus citri* (Brown & Nelson-Rees 1961) and later confirmed with expression patterns of genetic markers (Brown & Nur 1964; Brown & Wiegmann 1969; Brown 1972). Explicit demonstration of inhibition of RNA production from the heterochromatic set was further provided with autoradiographic essays (Berlowitz 1965). Males are thus seemingly transcriptional haploids; however, extensive evidence suggests that paternal chromosomes retain residual activity. First, irradiation of fathers does not reduce viability of their male offspring but yields them sterile, implying involvement of paternal chromosomes in reproductive function (Nelson-Rees 1962). Second, male offspring of hybrid crosses are often inviable (Nur & Chandra 1963; Rotundo & Tremblay 1982; Tranfaglia & Tremblay 1982; Kol-Maimon *et al.* 2014a; Chapter 2), even though they should escape hybrid incompatibilities if expression was strictly limited to maternally-inherited chromosomes. Third, and more conclusively, condensation of paternal chromosomes is reversed in certain tissues in *P. citri* and other mealybug species, including some somatic parts of the testes (Nur 1966; Nur 1967). Despite these findings, no direct evidence of residual transcriptional activity of paternal

genome in mealybugs—or other PGE taxa with somatic condensation—is available to date.

RNA-sequencing-based allele-specific expression analysis (ASE) is a direct approach to address this question, as it allows detecting parent-of-origin effects on gene expression at a whole-genome scale (Wang & Clark 2014). The rationale behind this strategy relies on estimating the proportion of transcripts that are generated from each parental allele in the F1 offspring of divergent strains or species by quantifying the ratio of maternal to paternal reads at discriminant SNP positions between the parental genomes. This approach has been extensively used to detect imprinted genes that are expressed only when maternally or paternally-inherited in diploid taxa, mostly vertebrates—e.g., mouse (Babak *et al.* 2008; Wang *et al.* 2008; Gregg *et al.* 2010; Wang *et al.* 2011; DeVeale *et al.* 2012), human (Heap *et al.* 2009; Metsalu *et al.* 2014; GTEx Consortium 2015; Hamada *et al.* 2016; Santoni *et al.* 2017), cow (Chamberlain *et al.* 2015; Chen *et al.* 2016), chicken (Zhuo *et al.* 2017)—and plants (Waters *et al.* 2011; Gehring *et al.* 2011; Luo *et al.* 2011). Recently, ASE studies have been extended to a wider taxonomic range, including invertebrate species with alternative genetic systems—e.g. the arrhenotokous honeybees and *Nasonia* wasps (Kocher *et al.* 2015; Wang *et al.* 2016)—and other taxa of evolutionary interest such as Schistosomatidae parasites, trematodes which have evolved separate sexes from an ancestral hermaphrodite system shared by all other trematode families (Kincaid-Smith *et al.* 2018). Also, from an evolutionary perspective, ASE studies have been employed to validate theoretical predictions on imprinting and genetic conflict (Babak *et al.* 2008; Wilkins *et al.* 2016).

In this chapter, I analysed ASE patterns in hybrid male offspring of *P. citri* and the closely related *P. ficus* to determine whether, and if so to what extent, paternal genomes are expressed in mealybugs. To do so, I called and validated species-specific SNPs from parental genome sequences and generated transcriptome data from somatic and reproductive tissues of adult F1 hybrid males (♀*P. citri* x ♂*P. ficus*). Under full transcriptional shutdown of paternal chromosomes, only maternal bases are expected to be found at informative SNP positions in F1 transcriptomes. Therefore, expression of paternal alleles can be revealed by incomplete bias to the maternal genome in all or a fraction of expressed genes with informative SNPs. I analysed soma and testes separately to compare patterns between these tissues and test two predictions. First,

that silencing is not complete in both soma and testis, as suggested by cytological and hybrid viability data; second, that ASE patterns differ between soma and testis. The latter is expected under an evolutionary arms race, and can result in two possible outcomes: higher levels of paternal allele expression (revealing ongoing resistance against elimination) or tighter silencing (due to successful maternal genome countermoves to assure complete incorporation into sperm).

I obtained ASE data for >7,000 genes in soma and testis and estimated that less than a third of soma-expressed and 80% of testis-expressed genes exhibit complete maternal monoallelic expression. While most of the genes that are not exclusively maternally-expressed still show a global bias towards the maternal genome, I was also able to identify 49 soma and 6 testis genes with biparental expression and 42 soma and 12 testis genes that are preferentially or exclusively expressed from paternal alleles. Most of these genes are involved in mitochondrial energy production and lipid metabolism, suggesting that paternal chromosomes are reactivated to boost transcription of genes involved in these key processes. Together, these results conclusively show that paternal alleles are expressed in mealybug males, most likely in a tissue-specific manner following decondensation of the heterochromatic paternal chromosomes. This study offers a solid foundation for further work aimed at characterising tissue- or stage-specific patterns of paternal genome activity and investigating putative paternal adaptations to resist maternal control under this bizarre genetic system.

3.3. Materials and methods

3.3.1. Experimental populations and hybrid crosses

Hybrid crosses were conducted between individuals from wild-derived, highly inbred laboratory strains: WYE3-2 (*P. citri*, >25 generations of sib-mating) and PF1-1 (*P. ficus*, >10 generations of sib-mating). Males and females from both parental lines were isolated before sexual maturity and isolated until adulthood. Hybrid crosses were set by placing pools (10-20 individuals) of brothers from the paternal species and sisters from

the maternal species in in 6cm-diameter glass Petri dishes. To encourage mating, a filter paper impregnated with 10 ng of synthetic sex pheromone from the paternal species was placed in the Petri dish. Geographical origin of WYE3-2 and PF1-1, rearing methods and composition of the synthetic sex pheromone dilutions were as described in Chapter 2.

After all males in the Petri dishes died, females were transferred to rearing bottles to lay eggs. Hybrid F1 offspring were reared until becoming sexually differentiated (third instar). At that stage, males were transferred to glass vials to reach adulthood. Due to extreme male specific mortality in hybrid offspring of crosses between *P. ficus* females and *P. citri* males, I was unable to obtain viable male hybrids from these crosses. Therefore, only hybrid males from *P. citri* mothers and *P. ficus* fathers could be sequenced.

3.3.2. RNA and gDNA extraction and sequencing

RNA was extracted from somatic tissues and testes of three pools of >70 adult F1 males (R1, R3 and R6). Testis dissections were performed in RNAlater (Thermo Fisher Scientific, USA). Soma and testes were immediately transferred to ice-cold TRIzol (Invitrogen) and stored at -80°C. RNA from soma (S1, S3 and S6) and testis replicates (T1, T3, T6) was extracted using isopropanol and chloroform (2.5:1) and linear acrylamide as a carrier. After extraction, residual gDNA digestion was performed using DNAse I (Thermo Fisher Scientific, USA) and RNA samples were purified with RNA Clean & Concentrator™-5 (Zymo Research, USA) and validated using the Bioanalyzer RNA 6000 Nano kit (Agilent).

Due to low RNA yield, cDNA amplification was performed using the Ovation® RNA-Seq System V2 (NuGen, USA). Two independent cDNA amplifications from each soma and tissue samples were performed separately on each sample to be sequenced as technical replicates. cDNA samples were purified using MinElute® Reaction Cleanup Kit (QIAGEN, The Netherlands) in TE buffer. DNA concentration was measured using Qubit® dsDNA BR Assay Kit and submitted for quality control and sequencing at Edinburgh Genomics (UK). 12 TruSeq Nano libraries with 350 bp insert size (three biological replicates of soma and testis, with two technical replicates of each) were

prepared and sequenced on two lanes on the Illumina HiSeq 4000 instrument (75 bp paired-end reads).

In addition to hybrid transcriptomes, I generated three biological replicates of RNA-seq data from pools of whole adult *P. citri* males in triplicate. Three TruSeq stranded mRNA-seq libraries were prepared from total RNA, one of which was sequenced on the Illumina HiSeq 4000 instrument (75 bp paired-end reads). The remaining libraries were sequenced on two lanes in the Illumina NovaSeq S2 instrument (50 bp paired-end reads).

Genomic DNA was extracted from 5-10 adult *P. citri* and *P. ficus* virgin females. Sample lysis, proteinase K digestion and RNA removal were performed using a DNeasy Blood & Tissue kit (Qiagen, The Netherlands), isolation of gDNA was carried out with a Wizard Genomic DNA Purification Kit (Promega, USA) according to manufacturer's instructions and Illumina sequencing was performed at Edinburgh Genomics on the HiSeq 2500 (from TruSeq libraries with 350 bp insert size).

3.3.3. Genome assembly, gene prediction and SNP calling

The genome assemblies obtained for the *P. citri* and *P. ficus* used in this study, PCITRI.V0 and PFICUS.V0, are deposited in http://www.mealybug.org. Gene prediction of both PCITRI.V0 and PFICUS.V0 assemblies was performed with the BRAKER1 pipeline (Hoff *et al.* 2015) using all RNA-seq reads generated from soma and testis replicates (see below for RNA-seq mapping). BUSCO v2.0.1 (Simão *et al.* 2015) was used to assess completeness of genome annotations and RepeatMasker v4.0.6 (http://www.repeatmasker.org) was used to screen the genomes for interspersed repeats and low complexity DNA sequences. A summary of assembly and annotation statistics is provided in Table S1 of Appendix 3.

To identify informative species-specific SNP positions allowing parent-of-origin assignments of reads, raw *P. ficus* genomic reads were mapped to the PCITRI.v0 assembly with bwa 0.7.15-r1140 (BWA-MEM algorithm) (Li 2013) and variants were called using FreeBayes v1.1.0 (Garrison & Marth 2012) with the following settings: --haplotype-length 0 --standard-filters --min-alternate-count 30 --min-alternate-fraction 0.9

--min-coverage 10 --use-best-n-alleles 2 -p 2 -J --pooled-discrete --pooled-continuous. A set of strict criteria was then applied to this raw set of discriminating variants to discard likely false positives supported by reads on single strands or mapping to one side of candidate variants only and remove non-SNP positions and polymorphic sites in *P. ficus* reads. After manual inspection of F1 transcriptome alignments, a further refined final set of discriminant ASE sites was obtained by removing SNPs that were not completely monomorphic for the reference PCITRI.V0 allele within *P. citri* transcriptomes. To do so, polymorphic sites and non-reference monomorphic sites within *P. citri* were called by remapping *P. citri* HiSeq genomic reads to PCITRI.V0 and running FreeBayes with the following settings: --haplotype-length 0 --min-alternate-fraction 0.05 --pooled-continuous (i.e. no base or mapping quality requirements).

3.3.4. RNA-seq mapping and parental allele counts and validation of ASE sites

An average of 24.97±0.83 million reads for soma and 26.39±0.77 million reads for testis were obtained from each sequencing library in a single lane. Initial quality was performed with FastQC v0.11.5 and reads were trimmed using Trimmomatic v0.36 (Bolger *et al.* 2014) with default settings. For all soma and tissue samples, paired reads for each lane and technical replicate were separately mapped to the reference genomes from both species using STAR v2.5.2b (Dobin *et al.* 2013) in the two-pass mode. Respectively, 61-68% and 47-52% of reads in each sample mapped uniquely to PCITRI.V0 and PFICUS.V0, with high consistency between lanes and technical replicates (Appendix 3, Fig S1). RSEM v1.3.0 (Li & Dewey 2011) was used to quantify expression as TPM (transcript per million) across samples.

To obtain reference (maternal = *P. citri*) and alternate (paternal = *P. ficus*) read counts at ASE sites, a single BAM file was generated for each biological replicate by merging data from both lanes and technical replicates. The ASEReadCounter walker from GATK v3.7 (McKenna *et al.* 2010) was used with the following settings: -U ALLOW_N_CIGAR_READS --minDepth 30 --minMappingQuality 30 --minBaseQuality 20.

A series of filters were then applied to remove problematic ASE sites within F1 transcriptomes. First, sites uniquely supported by one or two soma or tissue F1 replicates

were filtered out, thus keeping only ASE sites supported by data in all three tissue replicates. Second, ASE sites with >10% reads supporting bases other than reference and alternate were removed. At this point, manual validation of ~100 ASE sites showing intermediate degrees of paternal and maternal expression was performed by manually inspecting the alignments to the reference PCITRI.V0 genome of F1 transcriptomes and *P. citri* and *P. ficus* DNA-seq reads used for variant calling. In most of these ASE sites, expected genotypes (i.e. only alternate within *P. ficus* and only reference within *P. citri*) were confirmed, but a fraction of them were found to be monomorphic for the alternate base in *P. citri*, most likely due to errors in the reference genome or false negatives during variant calling within *P. citri*. Therefore, to remove false ASE sites from the final analysis, three RNA-seq replicates from whole adult *P. citri* males were mapped separately to PCITRI.V0 and counts for reference and alternate bases at all positions at ASE sites were obtained from the merged BAM file by running ASEReadCounter with default settings (i.e. no read depth or quality filters). Only informative sites with >95% of reference bases within the merged *P. citri* transcriptome were considered valid and kept in the F1 hybrid dataset. Lastly, for the final analysis I only included exonic ASE sites with a read depth of 30 in each F1 replicate for exons with at least two validated SNPs or an average read depth of 100 across F1 replicates for exons supported by a single SNP. SNP sites were assigned to genome annotation features using a custom script, asa2.

3.3.5. Parent-of-origin-specific gene expression analysis

To estimate to what extent gene expression patterns conform to expectations under lecanoid PGE (i.e. no expression of paternal alleles), parental-specific allele counts for all SNPs passing these filters were used to estimate their degree of bias towards the maternal genome. The method developed by Wang & Clark (2014) was broadly followed to calculate and test allelic expression differences, but adapted to an experimental setup lacking reciprocal F1 crosses. Bias to maternal genome was calculated as the proportion p_m of reference bases at each informative position, where p_m=0.5 denotes Mendelian expression (i.e. 1:1 ratio of maternal to paternal alleles) and p_m=1 corresponds to complete maternal expression. SNPs were divided into five

categories: maternal only ($p_m \geq 0.95$), maternally-biased ($0.65 < p_m < 0.95$), biparental ($0.35 \leq p_m \leq 0.65$), paternally-biased ($0.05 < p_m < 0.35$) and paternal only ($p_m \leq 0.05$).

For each gene, expression bias towards the maternal genome was calculated by pooling reference and alternate base counts across all coding regions. To statistically validate gene expression patterns and identify biparentally- and paternally-biased genes, exact binomial tests were conducted for each gene in all three F1 tissue replicates separately against the null hypothesis of Mendelian expression ($p_m=0.5$). To account for multiple testing, Bonferroni correction was applied and a P-value of 10^{-5} was considered as a cut-off for parentally-biased genes, so that genes over this threshold were considered to exhibit biparental expression. Also, a G-test of independence was performed for each gene to test whether expression biases were homogenous across all tissue replicates (FDR=0.01) (Wang *et al.* 2016). Genes that did not show significant heterogeneity across samples were immediately validated. The remaining genes were included in the final analysis only if all tissue replicates agreed on direction of bias and significance of exact binomial test. After removing the genes that failed to meet these criteria, combined expression bias to the maternal genome at gene level was determined by pooling paternal and maternal SNP counts from all replicates and performing a final exact binomial test for each gene with these pooled counts. Genes were considered maternally-biased when $P \leq 10^{-5}$ and $p_m > 0.5$ (fully maternally expressed when $p_m \geq 0.95$), paternally-biased with $P \leq 10^{-5}$ and $p_m < 0.5$ (exclusively paternal, $p_m \leq 0.05$) and biparentally expressed when $P > 10^{-5}$.

3.3.6. Functional annotation of genes

For all genes without complete maternal expression in soma and testis ($p_m < 0.95$), a GO enrichment analysis was performed to identify enriched functional terms within these genes escaping silencing compared to the remaining population of genes with a complete maternal expression pattern. A corrected FDR<0.1 was selected as a threshold to determine significant enrichment. To reduce bias of enrichment analysis (Timmons *et al.* 2015), the background gene population was restricted to genes with parent-of-origin information that passed the filters and were included in the ASE analysis.

Biparentally, paternally-biased and paternal-only genes were further investigated. To assign functions, associated GO terms and InterPro (IPR) domains were collected. Additionally, reciprocal orthologues in *Drosophila melanogaster* and *Acyrthosiphon pisum* were identified by comparing predicted proteins with the proteomes of *D. melanogaste*r using BLASTp v2.7.1+ (E-value≤$1e^{-25}$) and a modified version of the rbbh.py script (https://github.com/DRL). Tissue-specific expression patterns of D. melanogaster orthologues were obtained from FlyAtlas2 (Leader *et al.* 2017).

3.4. Results

To determine whether expression of paternally-inherited chromosomes is completely silenced in mealybug males, parent-of-origin-specific expression patterns were calculated in somatic and germline tissues of pooled F1 hybrid males originated from *P. citri* mother x *P. ficus* father crosses (Fig. 3.1). Transcriptomes from soma and testes were obtained from three biological replicates and mapped to the genome of the maternal species, *P. citri*.

3.4.1. Informative ASE sites

In order to assign parental origin of F1 transcriptomic reads, 4,533,219 fixed and discriminant SNPs (ASE sites) between *P. citri* and *P. ficus* were called by aligning DNA-seq reads to the PCITRI.V0 reference genome. Of these SNPs, 269,232 and 395,777 were found in at least one of the three biological replicates of F1 soma and testis transcriptomes, respectively, with >Q20 and a read depth of at least 30. Of these, 67% (179,212) and 64% (251,730) were present in all soma and testis replicates and thus considered for the ASE analysis.

As described in Materials and Methods, these ASE sites were surveyed to remove spurious or ambiguous positions and minimize the chances of inaccurate estimation of expression bias. A small fraction (0.15%) of these ASE sites (271 in

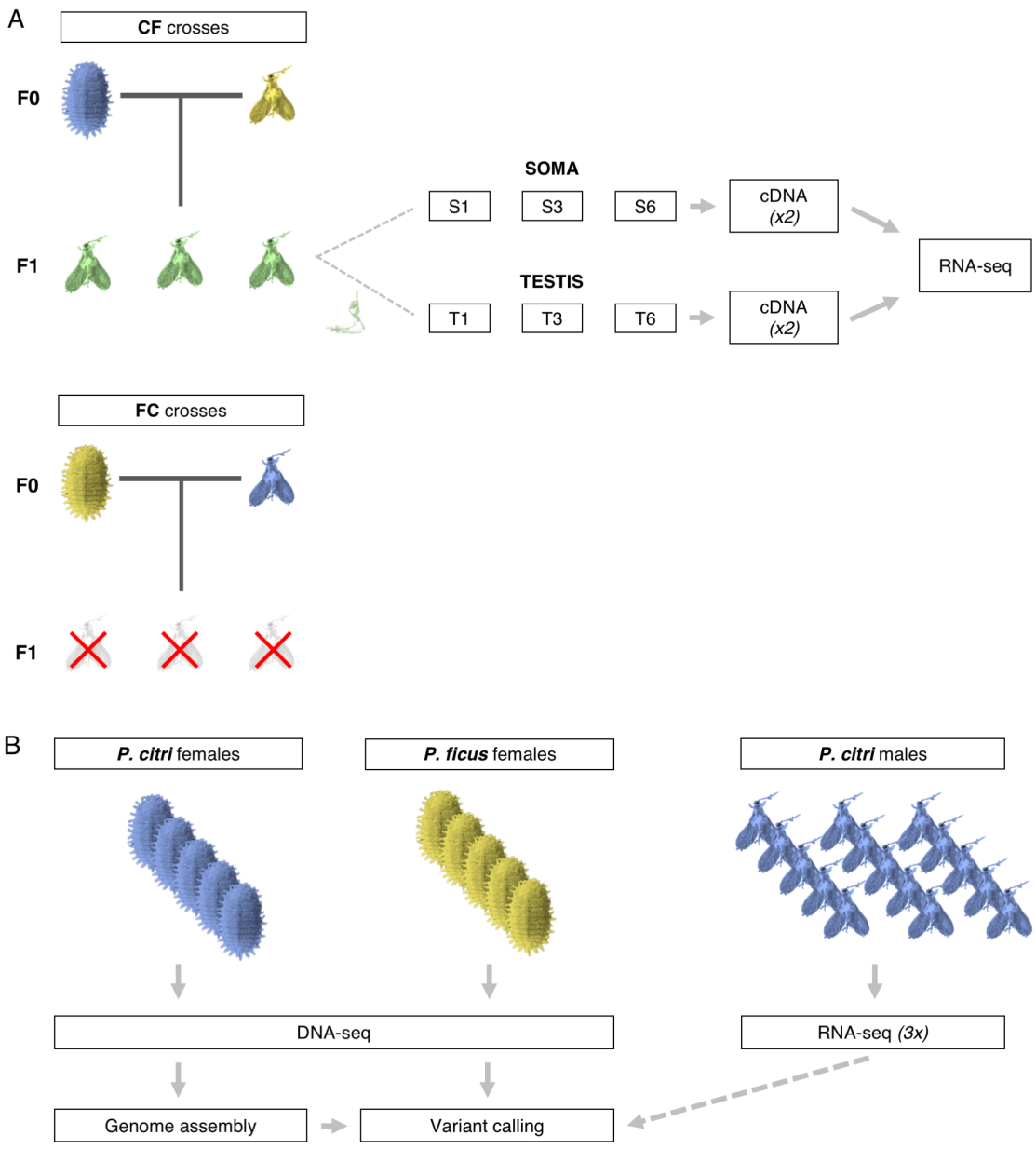


Figure 3.1. Experimental design. **(A)** Cross scheme between *P. citri* (yellow) and *P. ficus* (blue). Three mating pools from each reciprocal cross direction were set, of which only *P. citri*♀ x *P. ficus* ♂ crosses (CF) produced male offspring that survived to adulthood. Testes from F1 males (green) were dissected and sequenced separately from the soma. Two separate cDNA amplifications were performed and sequenced from each of the three soma and tissue F1 replicates. **(B)** To generate genome assemblies and call variants between *P. citri* and *P. ficus*, a pool of females from both species was sequenced. Additionally, to improve filtering of informative SNP positions, three transcriptomes from pools of adult *P. citri* males were obtained.

soma and 369 in testis) were removed due to the presence of bases other than reference and alternate in the transcriptome. Calling ASE in an additional *P. citri* male transcriptome allowed to identify an additional set of positions undetected during variant calling which contained non-reference bases within *P. citri* and could thus produce overestimation of paternal expression in hybrids. Out of 1,389,199 ASE sites found in the *P. citri* transcriptome (with no minimum quality or read depth requirements), 38,285 of these included more than 5% non-reference bases, of which 2,707 (in soma) and 3,511 (in testis) were found in the F1 transcriptomes and filtered out. Finally, to increase accuracy of genic ASE patterns, only exons with at least 2 SNPs or a single SNP with an average coverage of at least 100X between replicates were considered, which led to the removal of 5,318 ASE sites in soma and 4,657 in testis.

In total, 170,916 SNPs in soma and 243,193 in testis passed these filters (Appendix 3, Tables S2-S3). In soma, 94,758 (55%) of these ASE sites mapped to annotated exonic regions and were thus used to estimate parent-of-origin expression bias at gene level; in testis, the number of exonic ASE sites was 104,774 (43%).

3.4.2. Parent-of-origin-specific expression patterns at SNP and gene levels

For every informative ASE site in soma and testis F1 transcriptomes, maternal and paternal base counts were obtained to estimate bias to the maternal genome, p_m, at that position. In both soma and testis transcriptomes, expression was found to be strongly, but not completely, biased towards the maternal genome (Fig. 3.2). Only 41.6% of SNPs showed complete maternal expression (M, $p_m \geq 0.95$), while 54% were maternally-biased (MB, $0.65 < p_m < 0.95$), and a further 4% exhibited biparental expression (B, $0.35 \leq p_m \leq 0.65$). In testis, SNP expression patterns were closer to expectations under lecanoid PGE, as 87.6% are exclusively maternal (M); however, 11.8% are maternally-biased (MB) and less than 1% are completely biparental (B). Furthermore, a minority of SNPs (0.4% in soma and under 0.3% in testis) were preferentially (PB, $0.05 < p_m < 0.35$) or exclusively (P, $p_m \leq 0.05$) expressed from paternal alleles. For SNPs located within exons, expression profiles were globally similar but less biased to the maternal genome: SNPs showing incomplete maternal bias increased by a factor of 1.2-1.5 in both soma (to 65.4%) and testis (18.4%), while exclusively maternal SNPs dropped to 28.9% and 80.8%, respectively.

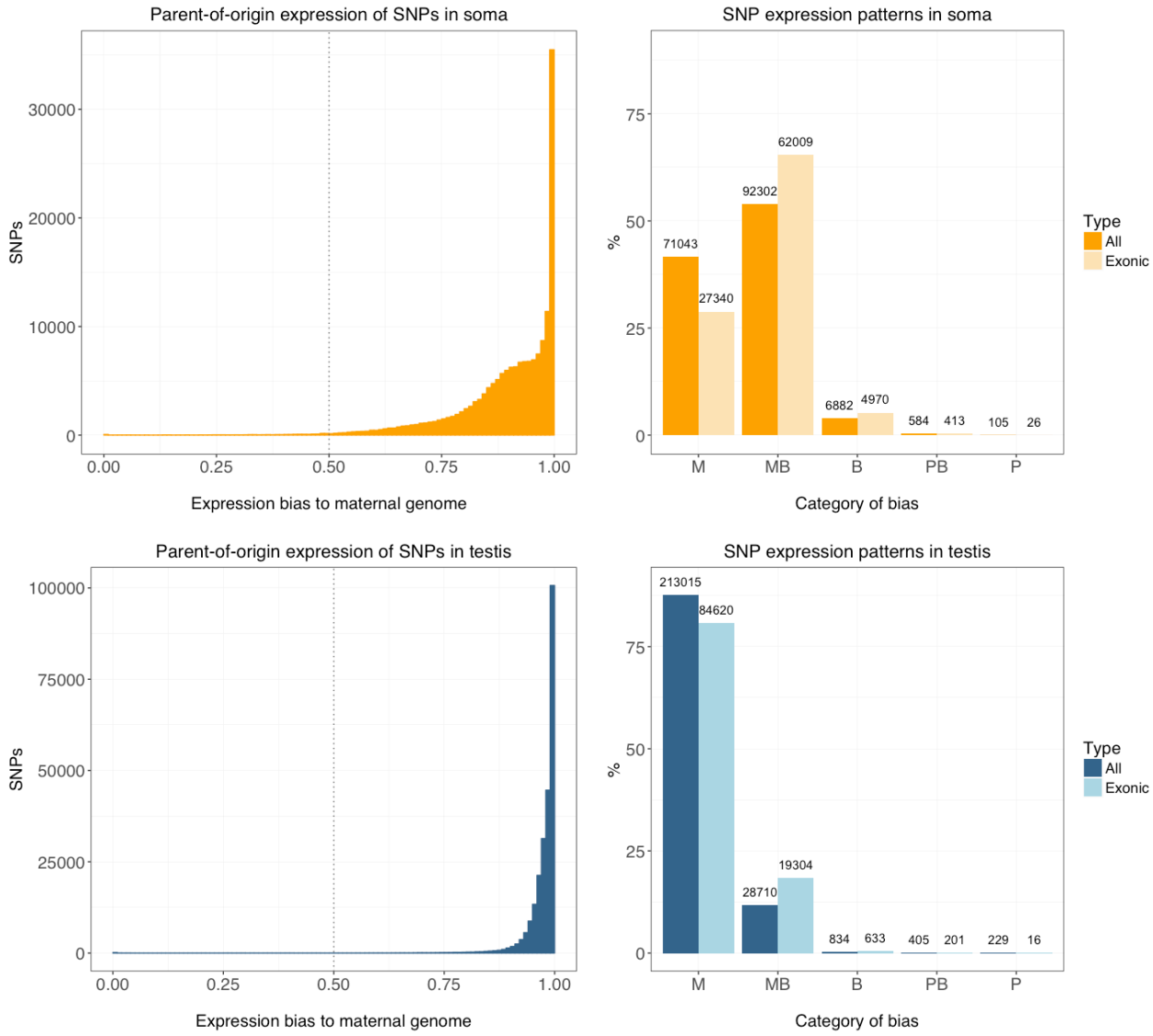


Figure 3.2. Quantification of parent-of-origin SNP expression patterns in soma (orange, top) and testis (blue, bottom) of F1 mealybug males. Left panels, histograms of expression biases to maternal genome (i.e. proportion of maternal bases at an ASE site) combined across replicates. Right panels, counts of total (in dark colours) and exonic (in light colours) SNPs partitioned by bias category.

Exonic SNPs were assigned to 6,368 genes with detectable expression levels (TPM>0) in soma and 6,737 in testis. Among these genes, 71 and 40, respectively, were discarded for showing high heterogeneity and discordance in bias between replicates. The remaining genes showed parent-of-origin expression patters analogous to those estimated at SNP level (Table 3.1). In soma, less than 30% of 6,287 genes were completely maternally-expressed (Fig. 3.3A), while in testis 80% belonged to this category (Fig. 3.3B). Most somatic genes (70%) and a fraction of testis-expressed genes (20%) were MB. In soma, 49 genes (less than 1%) were B, 40 were PB and 2 P. In testis, 6 were B, 10 PB and 2 P. Only moderate overlap was found between the genes showing biparental or paternal expression in soma and testis (Fig. 3.3C).

Table 3.1. Summary information and ASE patterns of genes with parent-of-origin specific information in soma and testis transcriptomes. M, maternal only. MB, maternally-biased. B, biparental. PB, paternally-biased. P, paternal only.

Genes with ASE information			
Tissue	**Total genes**	**Concordant genes**	**SNP density**
Soma	6,368	6,287	Per gene: 14.9 Per exon: 4.2
Testis	6,737	6,697	Per gene: 15.6 Per exon: 4.2

Category of expression bias					
Tissue	**M**	**MB**	**B**	**PB**	**P**
Soma	1,775 (28.2%)	4,421 (70.3%)	49 (0.8%)	40 (0.6%)	2 (<0.1%)
Testis	5,373 (80.2%)	1,306 (19.5%)	6 (0.1%)	10 (0.1%)	2 (<0.1%)

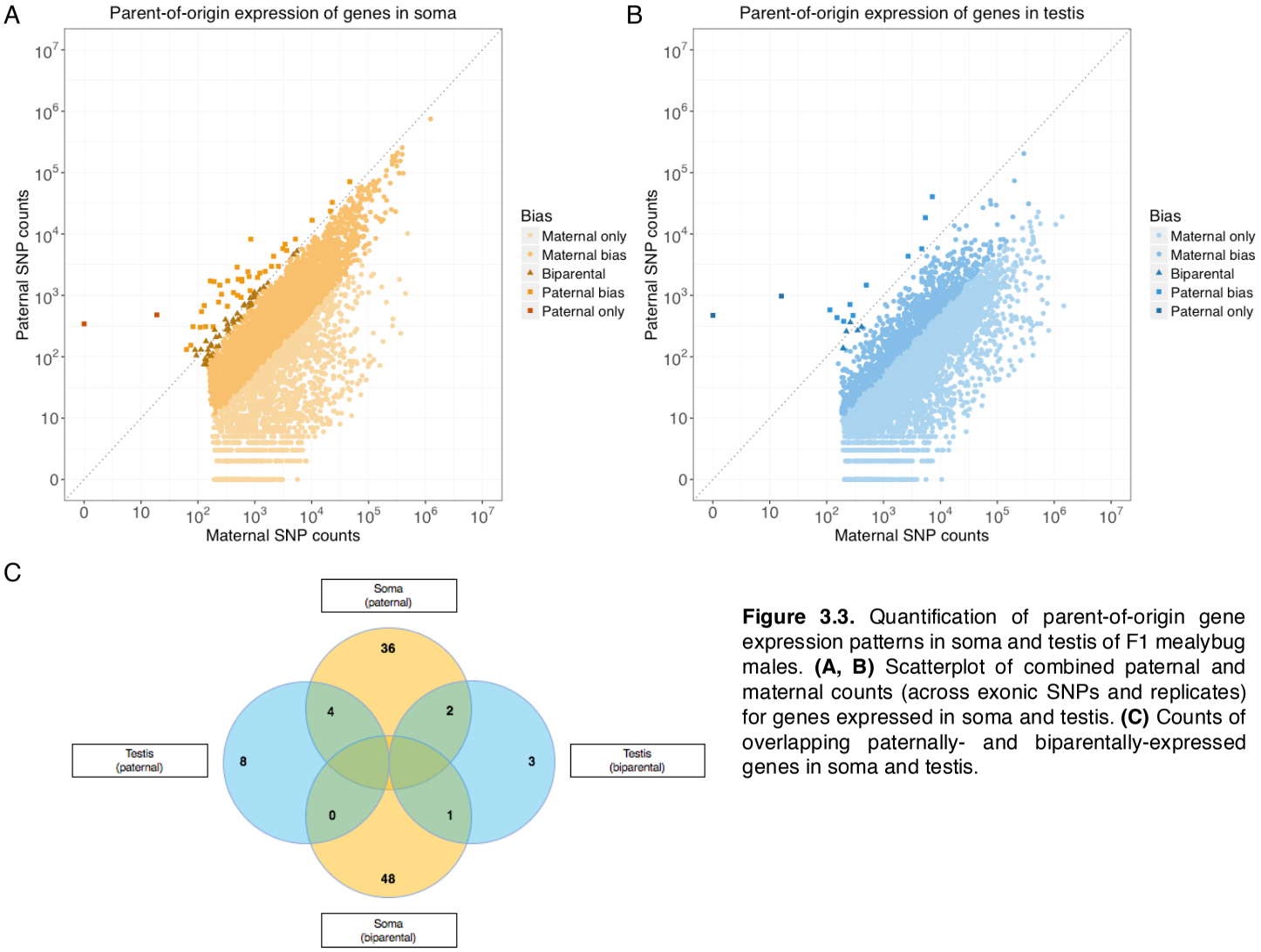


Figure 3.3. Quantification of parent-of-origin gene expression patterns in soma and testis of F1 mealybug males. **(A, B)** Scatterplot of combined paternal and maternal counts (across exonic SNPs and replicates) for genes expressed in soma and testis. **(C)** Counts of overlapping paternally- and biparentally-expressed genes in soma and testis.

3.4.3. GO enrichment analysis of genes without complete maternal expression

To obtain a functional profile of genes showing some degree of paternal expression (p_m<0.95) in mealybug hybrids, a GO enrichment analysis was performed in soma and testis separately. The analysis was applied to predicted gene products (5,083 in soma, 1,461 in testis) coded by genes with p_m<0.95 against all products coded by genes included in the ASA analysis (7,032 and 7,393). Next, the reduced set of genes with significant B, PB and P expression were more closely examined. To assign functions, I relied on a combination of GO terms, InterPro domains, reciprocal best BLASTp hits against *D. melanogaster* and *A. pisum* and additional BLAST searches for non-orthologs.

With respect to the whole set of genes with allele-specific expression information, genes without complete maternal expression in somatic tissues were enriched for the following biological processes (Fig 3.4): transcription (56 gene products), regulation of DNA-mediated transcription (116) and intracellular protein transport (35). Consistently, at cellular level, ribosome, nucleus and intracellular space were overrepresented; at molecular level, enriched terms were structural constituent of ribosomes and binding of nucleotides, ATP/GTP, DNA, RNA and metals. Putative functions were determined for 29/40 PB or P genes and 41/49 B genes (Appendix 3, Table S4). Most of these genes could be grouped in two categories: mitochondrial functions and fatty acid synthesis. Genes with mitochondrial expression included structural components of ribosome (3 small mitochondrial ribosomal subunit proteins, S29, S30 and S39-like, and 7 proteins from the large subunit: L4, L10, L19, L22, L28, L38 and L51, with expression patterns (p_m=0.33-0.51) ranging from PB to B), other ribosome-associated proteins (a GTPase and an additional genes with a rRNA-binding pentatricopeptide domain), components of the electron transport chain (ubiB, ubiE and CIA30 homologues) and proteins involved in mitochondrial fatty acid synthesis and oxidation. In addition to these last genes, 10 putative fatty acid synthases were identified, 9 of which were PB (p_m=0.14-0.35) and one B (p_m=0.5). Other major roles included transmembrane proteins (two ABC and three MFS transporters, and two ligand-gated ion channel-containing receptors), glucose metabolism (phosphofructokinase, glyceraldehyde-3-phosphate dehydrogenase, inositol-3-phosphate synthase, UDP

glycosyltransferase and ribulose-phosphate-3-epimerase), extracellular proteins (two laminin and a prophenoloxidase orthologs) and two putative retrotransposons with strongly PB expression (p_m=0.09-0.14).

Figure 3.4. GO enrichment analysis of genes without complete maternal expression (i.e. proteins coded by genes with $p_m < 0.95$) in soma. Enrichment scores for each GO term are represented as the –log10 of corrected P-values (using Benjamini–Hochberg false discovery rate). GO terms are grouped by classes (green, biological process; red, cellular component; purple, molecular function). Significantly enriched GO terms are indicated by point-up triangles (point-down for underrepresented GO terms). Triangle sizes are proportional to number of proteins associated with each GO term.

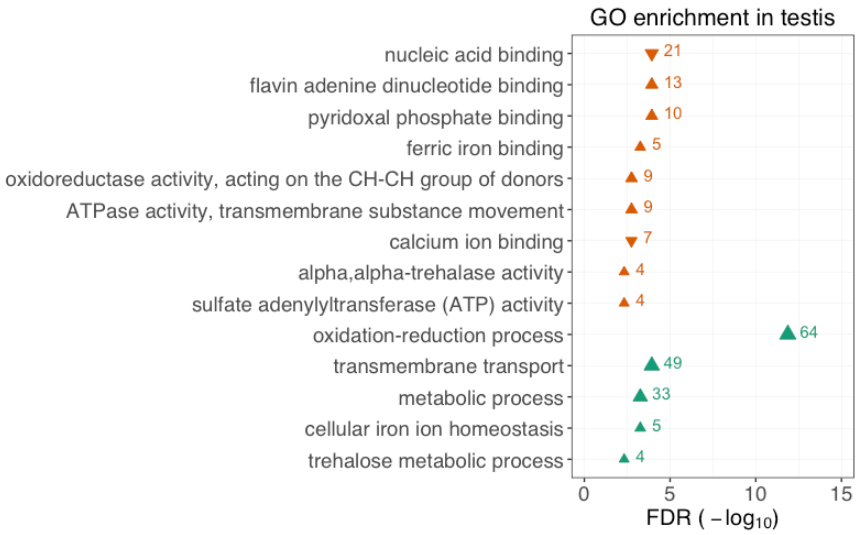


Figure 3.5. GO enrichment analysis of genes without complete maternal expression ($p_m < 0.95$) in testes. Colours and triangles as in Fig. 3.4.

In testis (Fig 3.5), the following GO biological processes were enriched in genes showing some degree of paternal expression: oxidation-reduction (64 transcripts), transmembrane transport (49), cellular iron homeostasis (5), metabolic processes (33) and trehalose metabolism (4). Functional roles could be determined for 11/18 P, PB and B genes (Appendix 3, Table S5). In contrast to somatic genes, mitochondrial functions were only found in a single gene, orthologue to *Air* in *D. melanogaster*, and only two genes associated to fatty acid metabolism were identified.

3.5. Discussion

Whole-genome meiotic drive of maternally-inherited chromosomes is common to all 7 independent origins of PGE described so far in arthropods, yet silencing of paternal allele expression in male soma is an additional manifestation of this genetic system in at least 3 of those origins (Burt & Trivers 2006; Gardner & Ross 2014; Hodson *et al.* 2017). In mealybugs, paternal chromosomes are heterochromatised (thus presumably transcriptionally silenced) during development and elimination is delayed until spermatogenesis. Therefore, males fail to express paternally-inherited traits, such as dominant lethal chromosomal derangements in mealybugs (Brown & Nelson-Rees 1961) or mutations affecting eye colour or wing morphology (Brown & Wiegmann 1969).

However, it has long been suspected that paternal chromosomes must retain some residual activity in mealybugs, since heterochromatinization does not completely protect from deleterious mutations in the paternal genome or hybrid incompatibilities (Nelson-Rees 1962; Nur & Chandra 1963; Brown & Nur 1964; Rotundo & Tremblay 1982; Tranfaglia & Tremblay 1982; Kol-Maimon *et al.* 2014a). To date, the most convincing—albeit indirect—evidence of expression of paternal alleles in *P. citri* and other mealybug species was the finding that heterochromatinization of paternal chromosomes is reversed in cells of certain male tissues in a stage- and species-dependent manner, likely indicating paternal genome reactivation (Nur 1966; Nur 1967; Nur 1970). In this chapter, I explored patterns of parent-of-origin specific gene expression in hybrid male mealybugs to obtain direct proof of expression of paternal alleles under this form of PGE.

3.5.1. Paternal alleles are expressed in a tissue-specific manner

Expression of paternal alleles was found in both soma and testis, yet to a different degree. In somatic tissues, less than a third of all genes containing informative SNPs and passing quality filters showed complete maternal expression, showing that reactivation of paternal alleles affects most of the genome and is not restricted to a reduced number of genes. The most convincing explanation for this phenomenon is—as first suggested by Nur (1967)—loss of heterochromatinization across a range of somatic cells triggering reactivation of the transcriptional activity of paternal chromosomes. In mealybugs, it seems unlikely that paternal alleles can be transcribed when in a heterochromatic state, since there is direct evidence of RNA synthesis inhibition from the heterochromatic set (Berlowitz 1965). Also, the best known example of whole chromosomal silencing—inactivation of one of the X chromosomes for dosage compensation in mammals—is also achieved via heterochromatinization (Lyon 1961; Wutz 2011). A relatively small proportion of genes (1-3% in mouse, 15% in human) on inactive X chromosomes in females are known to escape inactivation to some degree (Carrel & Willard 2005; Yang *et al.* 2010; Crowley *et al.* 2015). Although little is known of how these genes escape inactivation, they seem are able to do so in a gene-specific manner via the recruitment of chromatin regulators (Heard & Disteche 2006; Filippova *et al.* 2005).

In mealybugs, however, the large number of genes escaping complete paternal silencing to some degree (up to 70% in soma) suggests a likely reversal of the heterochromatic status of the whole paternal set, rather than a gene-by-gene reactivation mechanism in chromosomes otherwise maintaining their condensed status. In addition to cytological observations of reversal, in which decondensed paternal chromatin becomes indistinguishable from the maternal (Nur 1966; Nur 1967), this view is supported by detection of incomplete bias to the maternal genome in intergenic regions in the soma (p_m=0.93 on average across the whole genome, Table S3 in Appendix 3). Although the functional significance of extragenic transcription is still debated (Palazzo & Lee 2015), the data indicates that paternally-inherited chromosomes contribute modestly to this process in mealybug males. Reversal of heterochromatinization in adults is thus not specific to coding regions, yet tissue-specific patterns of reactivation of paternal chromosomes remain an open question. This analysis revealed a strong difference in patterns of bias to the maternal genome between soma and testis, where only 20% of genes show some degree of paternal expression. This difference indicates that there is not a unique global pattern of reactivation across the male and that individual tissues show variability in the extent to which paternal chromosomes are expressed. Since RNA was extracted from whole males (minus the testis), it is not possible to profile allele-specific expression variation across somatic tissues, so that any inference in this regard must be drawn from functional annotation of B or PB genes.

3.5.2. Sequencing of adult tissues might not capture putative paternal adaptations acting on spermatogenesis

Testes were specifically targeted in this study as they are the arena of elimination of paternal chromosomes under germline PGE. Therefore, reactivation of paternal alleles in germline tissues might represent a paternal response to fight their exclusion from sperm (Herrick & Seger 1999; Ross *et al.* 2010a). In *P. citri*, paternal chromosomes in cells of the cyst wall of the testes lose their heterochromatic status in early larval stages (Nur 1967). Germline cysts seem to be involved in packaging of sperm into bundles and, potentially, establishment of epigenetic marks (Buglia & Ferraro 2004). If expression of paternal alleles in these cells could interfere with PGE is unclear. Under the conflict hypothesis, and depending on the current dynamics of a putative arms race between

parental alleles over spermatogenesis, testes could be expected to be either a hotspot for paternal chromosome reactivation (paternal adaptations) or a tissue in which the control exerted by the maternal genome over their paternal counterpart is more severe (maternal countermoves). The stark contrast with patterns of allele-specific expression in the rest of the males reveals that paternal chromosomes are more tightly silenced in testes, seemingly supporting the latter scenario. However, it must be noted that male meiosis and the subsequent degradation of spermatids containing paternal chromosomes mostly takes place prior to adulthood, during second and third larval instars (Nur 1962b; Bongiorni *et al.* 2004). RNA-seq of adult male testis could then miss paternally-expressed anti-PGE responses, and the higher maternal bias in testes could derive from transcriptomic activity in sperm, which only contain maternally-derived chromosomes. Our annotation of the few biparentally and paternally-expressed genes and analysis of enriched or depleted GO terms among incompletely maternally-expressed genes did not identify any genes specifically involved in reproductive functions in adult testes. Therefore, although these results are inconclusive on whether there is scope for paternal adaptations to escape elimination, the demonstration that paternal alleles are both also expressed in testis yet apparently submitted to a stricter control can be a first ground to further explore this hypothesis. The pipeline developed in this chapter would be a valuable approach to investigate possible fluctuations in bias to the maternal genome in testis-expressed genes across spermatogenesis progression in younger males.

3.5.3. Reactivation of paternal chromosomes could represent and adaptation to increase transcription in certain tissues

The patterns of parent-of-origin gene expression found in this study raises the question of whether transcriptional reactivation of paternal chromosomes is caused by failure of paternal chromosome silencing, either caused by paternal resistance or by intrinsic errors of the silencing machinery, or has an alternative explanation. Paternal chromosome heterochromatinization takes place during the 7th cleavage division of male embryos, most likely triggered by maternally-deposited factors in the cytoplasm (Prantera & Bongiorni 2012). After this point, the maintenance of heterochromatic state depends directly on the maternally-inherited set (Nur 1962b; Chandra 1963; Brown & Nur 1964; Ross *et al.* 2010a). Reproducible temporal and interspecific variation in tissues

that undergo reversal suggest that decondensation of paternal chromosomes is a controlled process. Furthermore, hybrid male offspring of crosses between different *Pseudoccocus* species only undergo (occasionally incomplete) reversal in the same tissues as in males from the maternal species (Nur 1967; Nur 1972). Reversal, then, seems to be maternally controlled, either directly (by cytoplasmic factors) or indirectly (by the maternally-inherited chromosomes) and therefore in the interest of the maternal genome.

Genes with some degree of somatic paternal expression are enriched in functional categories associated with transcription and translation. Moreover, the majority of B and PB genes with assigned functions are involved in energy production and fatty acid metabolism (Tables S4-S5 in Appendix 3). In soma, these genes encode at least 10 mitochondrial ribosome proteins, in addition to others involved in the mitochondrial electron transfer chain and cytoplasmic ATP and NADH production. Mitochondrial ribosome proteins (MRPs) are encoded by nuclear genes and imported into the mitochondria, where they are assembled into ribosomal subunits in conjunction with mitochondrially-encoded rRNAs (Tselykh *et al.* 2005; Richman *et al.* 2014; Rackham & Filipovska 2014). Mitochondrial ribosomes are responsible for translation of mRNAs encoded by mitochondrial genes, including those involved in ATP production (O'Brien 2002; Beckmann & Herrmann 2015). A possible reason for the overrepresentation of MRPs genes might be their accelerated evolution rates (O'Brien 2003; Desmond *et al.* 2011), which allows a more reliable estimation of allele-specific expression due to higher SNP density. Among genes involved in fatty acid metabolism, those coding for fatty acid synthase (FAS) and enzymes acting on CoA thioesters were predominant. FAS are cytoplasmic multifunctional proteins responsible for *de novo* biosynthesis of the fatty acid palmitate from CoA thioester precursors (Smith *et al.* 2003). Palmitate and its derivatives are involved in several key process, including energy production, protein modification and localization, signalling, diapause, stress tolerance and cell membrane composition (Stanley-Samuelson *et al.* 1988; Jones & Infante 2015). Mealybug males cease to feed at late second instar (Franco *et al.* 2009), so upregulating transcription of FAS and other lipid related genes involved in synthesis and metabolism of fatty acids could be a mechanism to compensate for dietary restriction. Interestingly, one of the cell types that can undergo heterochromatization reversal in *P. citri* are oenocytes (Nur 1966), which

are specialised in lipid metabolism (Makki *et al.* 2014). From here, it can be suggested that paternal chromosomes are recruited to upscale energy production and lipid metabolism, among other processes, in a tissue-dependent fashion.

Considering the functional profiles of biparentally and paternally-biased genes found in the soma in this study, I conclude that reversal of heterochromatinization is an adaptive phenomenon to boost transcription to diploid levels in a tissue-dependent manner. The notion that paternal chromosomes are reactivated to increase transcription had been already tentatively suggested by Nur (1966), who also discovered an alternative—and possibly complementary—mechanism to increase transcription in males: endomitosis of euchromatic (i.e. decondensed) chromosomes (Nur 1968). The relationship between reversal and endopolyplodization is very suggestive: only euchromatic chromosomes undergo endomitosis and reversal tends to be observed in tissues that later increase in ploidy, such as oenocytes and Malpighian tubules (Nur 1967; Nur 1968; Lorick 1970). To date, it was thought that the main mechanism for males to increase genetic activity was by having more cells in most tissues—yet, tellingly, not in Malpighian tubules, where cell numbers are equivalent between males and females (Berlowitz *et al.* 1968). Reversal of heterochromatic chromosomes, both on its own and as a prerequisite for endopolyploidization, would be a less costly way to increase transcriptional activity in tissues with high metabolic demands.

3.5.4. Conclusion and future directions

In this chapter, I have conclusively demonstrated a strong parent-of-origin effect in gene expression in lecanoid PGE, whereby maternal chromosomes are the main contributors to transcriptomic activity in males due to heterochromatization of the paternal set. To my knowledge, this analysis constitutes the first study in which a next-generation sequencing approach has been employed to reveal such effects in a PGE species. These patterns of bias to the maternal genome are also highly consistent between replicates. However, regulation of gene expression in males is not as simple as complete silencing of paternal chromosomes, as paternal alleles are also transcribed to some degree across most the genome, most likely to meet tissue-specific transcription requirements. This finding has important implications for our understanding of PGE, since the assumption that elimination and heterochromatinization of paternal

chromosomes under PGE have an identical result—complete haploid expression of maternally-inherited chromosomes—is common in the literature (see e.g. Brown 1964; Sager & Kitchin 1975; Brun *et al.* 1995; Normark 2004; de la Filia *et al.* 2015). In addition to determining genome-wide levels of paternal expression, I have identified a number of genes with biparental and paternally-biased expression which will be candidates for downstream studies to understand functional consequences of PGE. However, it must be noted that the aim of this study was not to obtain a comprehensive list of biparentally-expressed genes, but to evaluate whether paternal chromosomes are completely silenced across the genome using a conservative approach. Stricter SNP validation criteria and allele-specific bias categorization than in other allele-specific studies (Wang & Clark 2014) inevitably led to loss of power to detect additional genes (specially with tissue-expression enrichment), but future studies using tissue-specific sequencing or a more relaxed filtering criteria could result in an expanded list.

From an intragenomic conflict perspective, the differential parent-of-origin-specific expression patterns between somatic and reproductive tissues of mealybug males found in this study suggest an coevolutionary scenario over maintenance of paternal silencing in which the maternal genome appears to have strengthened control of paternal alleles in testis. However, there are other potential explanations for these patterns that cannot be fully excluded. Targeting earlier larval stages, as discussed earlier, would allow capturing spermatogenesis and framing paternal genome expression patterns in the context of their elimination. Finally, another exciting prospect for the analysis strategy developed in this chapter is its application to other germline PGE taxa with and without paternal chromosome silencing, in order to gain a comparative understanding of the somatic manifestations of this genetic system and its evolutionary implications.

Chapter 4

The unusual reproductive system of head and body lice (*Pediculus humanus*)

As published:

De la Filia, A.G., Andrewes, S., Clark, J. M. & Ross, L., 2017. The unusual reproductive system of head and body lice (*Pediculus humanus*). *Medical and Veterinary Entomology*, 32, pp. 226–234.

4.1. Summary

Insect reproduction is extremely variable, but the implications of alternative genetic systems are often overlooked in studies on the evolution of insecticide resistance. Both ecotypes of *Pediculus humanus* (Phthiraptera: Pediculidae), the human head and body lice, are human ectoparasites, the control of which is challenged by the recent spread of resistance alleles. The present study conclusively establishes for the first time that both head and body lice reproduce through paternal genome elimination (PGE), an unusual genetic system in which males transmit only their maternally derived chromosomes. Here, we investigate inheritance patterns of parental genomes using a genotyping approach across families of both ecotypes and show that heterozygous males exclusively or preferentially pass on one allele only, whereas females transmit both in a Mendelian fashion. We do however observe occasional transmission of paternal chromosomes through males, representing the first known case of PGE in which whole-genome meiotic drive is incomplete. Finally, we discuss the potential implications of this finding for the evolution of resistance and invite the development of new theoretical models of how this knowledge might contribute to increasing the success of pediculicide-based management schemes.

4.2. Introduction

One of the most striking features of insects is the extraordinary diversity of their reproduction, which is unparalleled in any other animal group. This is illustrated by the wide heterogeneity of reproductive and genetic systems found across insect taxa that differ from the canonical diplodiploidy prevalent in metazoans (Normark 2003). The most well-known example of these alternative genetic systems is arguably arrhenotoky (i.e. haplodiploidy *sensu stricto*, whereby males develop from unfertilized eggs). However, many other more complex and bizarre non-diplodiploid systems have been described. Many of these are common in economically important insects: for instance, parthenogenesis (female reproduction without fertilization) is disproportionately

abundant in pest species, including representatives of groups such as mites, aphids and scale insects, compared with non-pest relatives (Hoffmann *et al.* 2008; Ross *et al.* 2013). Telling signs of alternative genetic systems are non-Mendelian inheritance patterns of traits or genetic markers, which are often discovered fortuitously in certain species but are rarely explored further despite their potential implications for key aspects of insect management, such as the evolution of virulence and insecticide resistance.

One of these species is the human louse, *Pediculus humanus*, a blood-sucking ectoparasite that occurs worldwide and causes infestations with serious medical, economic and social consequences. Human lice are divided into two ecotypes: the head louse (*Pediculus humanus capitis*) and the body louse (*Pediculus humanus humanus*) (Durden & Musser 1994), which differ in their ecology and clinical importance. Whereas body lice feed on human skin and lay eggs on clothes, head lice live and feed on the human scalp only. Epidemiologically, head louse infestations are more common and mostly affect children, regardless of economic status or geographic region (Clark *et al.* 2013). By contrast, body louse infestations are associated with lower socioeconomic conditions and pose a more serious health threat because the body louse is a vector of epidemic pathogenic bacteria, including *Bartonella quintana* (trench fever), *Borrelia recurrentis* (relapsing fever) and *Rickettsia prowazekii* (epidemic typhus) (Raoult & Roux 1999).

Control of human lice generally involves a combination of manual removal techniques and the use of diverse chemicals often referred to as pediculicides. However, many of the most widely used pediculicides have become ineffective as a result of the spread of resistant strains (see Durand *et al.* (2012) and references therein) and, as many pediculicides share common chemistry and targets (Clark *et al.* 2013), further spread of resistance is likely. To reduce this risk, it is important to unravel the molecular and metabolic mechanisms involved in pediculicide resistance (Oakeshott *et al.* 2003), which have been addressed by several studies in recent years (Yoon *et al.* 2008; Kwon *et al.* 2014). However, current understanding of how resistance evolves and spreads through populations is very limited because key factors such as population structure, gene flow, reproductive genetics, life history and mating system remain insufficiently explored. Better understanding of these factors and their roles in the evolution of pesticide resistance will support the development of successful novel treatment

strategies and management programmes aimed at preventing the spread of resistance genotypes.

Until recently, it was assumed that inheritance of traits such as pesticide resistance in lice would follow the classic laws of Mendelian genetics. However, an unexpected finding in the body louse suggested that whereas allele transmission in females followed Mendelian expectations, it was non-Mendelian in males: heterozygous male parents systematically passed on one of their two alleles to their offspring (McMeniman & Barker 2005). Moreover, the transmitted allele was of maternal origin in all cases and the paternally derived alternative allele was absent from the offspring. This mode of inheritance is consistent with paternal genome elimination (PGE), a type of haplodiploid reproduction found across several arthropod orders in which males do not transmit paternally inherited alleles to their offspring (Normark 2003). It is surprising that the possible presence of PGE in *P. humanus* has not been considered in the context of louse control because it may have implications for the evolution of pesticide resistance. Theoretical approaches have shown that haplodiploidy can accelerate the invasion of resistant alleles under certain circumstances (Crozier 1985; Caprio & Hoy 1995; Denholm *et al.* 1998; Carrière 2003) and PGE has been invoked to explain the rapid spread of insecticide resistance in New Caledonian populations of the coffee berry borer beetle *Hypothenemus hampei* (Brun *et al.* 1995). Furthermore, PGE is likely to elicit sex-specific responses and selection pressures that can further affect the way resistance genotypes spread through populations (Carrière 2003).

Although the study by McMeniman & Barker (2005) is suggestive of the presence of PGE in *P. humanus*, it requires further confirmation. They show that a proportion of heterozygous males transmit both alleles in a Mendelian fashion, which would mean that PGE was polymorphic in the study population (McMeniman & Barker 2005). This finding is unlike any form of PGE described so far, which has always been found to be complete. Further, McMeniman & Barker (2005) used only three markers in their study, which falls short of covering the whole genome and does not allow determination of whether drive is complete or restricted to some chromosomes. Moreover, the Culpepper strain (Culpepper 1944) used by McMeniman & Barker (2005) in their experiment might not be representative of natural populations as it has evolved under laboratory conditions since 1945 and has adapted to rabbit blood, rather than human. It is therefore possible that a

drive factor emerged in this strain independently of natural body louse populations, which were not sampled. Finally, the study by McMeniman & Barker (2005) included only body lice and no data on inheritance in head lice have been published since. Here, we study patterns of allele inheritance in both head and body louse families derived from recently collected natural populations reared on human blood.

In order to determine whether males show complete genome-wide meiotic drive consistent with PGE, we used a two-generation experimental crossing design and a panel of multiple polymorphic microsatellite markers. Transmission patterns were determined by genotyping both parents and their offspring to determine whether both alleles at a given heterozygous parental locus are present at a 50:50 ratio in the offspring (Mendelian transmission) or whether only one allele is transmitted by male parents (PGE). The current study provides the first reported evidence of PGE in the head louse and confirms its occurrence in body lice, albeit with no consistent evidence of a PGE polymorphism between males. We do, however, observe occasional leakage of paternal alleles, especially in body lice. Finally, we also suggest subsequent research directions aimed at increasing current understanding of how PGE operates in lice, particularly whether it affects gene expression patterns in males, and discuss the implications of this unusual genetic system for the evolution of parasitic lice in general and, most specifically, the evolution of pediculicide resistance.

4.3. Materials and methods

4.3.1. Experimental design

A series of intraspecific crosses were set up using individuals from the head louse strain SF-HL and the body louse strain Frisco-BL. The SF-HL colony was established in 2002 from head lice collected from ~20 infested children in Plantation, Miami and Homestead (FL, U.S.A.). Approximately 50 males and 50 females were used to temporarily establish a colony on human volunteers (Takano-Lee *et al.* 2003). Fertile eggs from Homestead were added to the colony at least three times between 2002 and 2006. Approximately 30–50 eggs were introduced each time. The sex ratio of the eggs

was assumed to be ~ 50:50. The colony was placed on an *in vitro* rearing system in 2006 (Yoon *et al.* 2006). The Frisco-BL colony of human body lice was originally collected from nine homeless individuals in San Francisco (CA, U.S.A.) by Dr Jane Koehler (University of California San Francisco Medical Center, San Francisco, CA, U.S.A.) in December 2008. Both colonies have been maintained by the Clark Laboratory at the University of Massachusetts-Amherst on human blood using the same *in vitro* rearing system (Yoon *et al.* 2006) under environmental conditions of 30 °C, 70% relative humidity and an LD 16:8 h photoperiod in rearing chambers (University of Massachusetts-Amherst Institutional Review Board approval no. E1404/001-002).

Parental generations (F_0) were established by random selection of pairs of sexually immature third instar lice from each colony. These pairs were transferred to individual cages. Lice were sexed after reaching reproductive maturity using the approach first described by Meinking (1999) and cages were checked for same-sex pairs. In these cases, a pair of male-only and female-only cages was selected at random and a randomly chosen individual was swapped between cages. After this point, all cages were screened daily to check for oviposition or the death of parents. Males were removed and stored in 100% ethanol at 4 °C after 7 days or immediately after their death. Females were allowed to lay eggs for 2 weeks and were then removed and stored in 100% ethanol at 4 °C. Offspring (F_1) of all crosses were raised until early third instar stage and then transferred to ethanol. In total, F_1 broods for 26 head and 13 body louse families were obtained.

4.3.2. DNA extraction and PCR amplification

Total genomic DNA from parents and body louse F_1 individuals was extracted with a DNeasy Blood and Tissue Extraction Kit (Qiagen Benelux BV, Venlo, the Netherlands). DNA from head louse F_1 individuals was extracted with a prepGEM Insect Kit (ZyGEM NZ Ltd, Hamilton, New Zealand) in a 20-µL reaction volume. A panel of three multiplexes (MX1, MX2 and MX4) from Ascunce *et al.* (2013) containing 12 microsatellite loci in total (T8_1, M3_10, M3_19, M2_2, T2_6, M2_19, M2_13, M2_3, T9_6, T2_7, T4_5 and T1_4) was used for polymerase chain reaction (PCR) amplification. The PCR reactions for each of the three multiplexes were performed using the Type-it Microsatellite PCR Kit (Qiagen Benelux BV) in a 15-µL reaction volume. Primer

sequences and reaction mixes were as described in supplementary Tables S1–3 in Ascunce *et al.* (2013). The PCR reactions were performed under the following conditions: initial denaturation at 95 °C for 5 min; 35 cycles of denaturation at 94 °C for 30 s; annealing at 52 °C for 45 s; extension at 72 °C for 45 s, and a final extension step at 72 °C for 40 min. One microlitre of PCR product for each sample and multiplex was sent to Edinburgh Genomics (University of Edinburgh) for genotyping on the ABI 3730 DNA Analyzer system (ThermoFisher Scientific, Inc., Waltham, MA, U.S.A.).

4.3.3. Primer mapping

To reveal the extent of the genome coverage of the microsatellite panel in use, all loci were mapped against the most recent publicly available louse genome assembly. All forward and reverse primer sequences were blasted against the U.S. Department of Agriculture strain genomic assembly (PhumU2) using the blast tool in VectorBase (National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, U.S.A.) with default settings.

4.3.4. Microsatellite scoring and data analysis

Upon reception of raw trace files, microsatellite alleles were scored using the Microsatellite Plugin implemented in Geneious version 8.1.3 (Biomatters Ltd, Auckland, New Zealand). Estimates of observed (H_O) and expected (H_E) heterozygosity, number of alleles and inbreeding coefficient F_{IS} (Weir & Cockerham 1984) per locus for F_0 populations were obtained using the online version of Genepop version 4.2 (Raymond & Rousset 1995; Rousset 2008) with default parameters. For each family and locus, paternal and maternal allele transmission ratios were calculated as the number of occurrences of one of the two alleles in the F_1 offspring divided by the total number of F_1 genotypes. Because of the clear expectation of allele transmission following McMeniman & Barker (2005) and other PGE species, these parental ratios were defined in different ways to represent these different sex-specific transmission patterns. For paternal transmission ratios, the allele used in this calculation was that with higher representation in the offspring genotypes. For maternal transmission ratios, one of the two alleles was chosen at random. Likewise, when both parents were heterozygous for

the same alleles at a given locus, parental allele counts were assigned under the assumption that the driving allele present in all offspring was paternally derived. Exact binomial tests to detect significant deviations from Mendelian expectations in each locus were performed in R version 3.2.4 (R Foundation for Statistical Computing, Vienna, Austria). To correct for multiple testing, a reduced significance level of α=0.01 is considered in addition to the conventional level of α=0.05.

4.4. Results

4.4.1. Informative parents and microsatellite panel

In order to determine patterns of allele transmission, the F_1 offspring of F_0 parents with at least one heterozygous locus were genotyped because parents that are homozygous for all loci are not informative. Multi-locus heterozygosity of parental populations was higher in head louse F_0 (H_O=0.351) than in body louse F_0 (H_O=0.256) despite higher allelic richness in the latter, as a result of the elevated inbreeding coefficient in the body louse population (F_{IS}>0.2) (Table 4.1). At least one heterozygous marker was found in all 26 head louse and 11 body louse fathers. Likewise, 24 head louse and all 13 body louse mothers were heterozygous for at least one locus. This allowed for the determination of both paternal and maternal allele transmission patterns in almost all families (Appendix 4, Table S1).

All F_0 and F_1 individuals were genotyped using the 12-locus microsatellite panel, but not all markers could be included in the analysis (Appendix 4, Tables S2 and S3).

Species	Families	F1/family	Loci	Allele/locus	HO	HE	FIS	Ho ♂	Ho ♀
Head louse	26	8-12	11 (10)	2.55±0.32	0.341±0.065	0.366±0.071	0.021±0.044	0.315±0.064	0.367±0.069
Body louse	13	7-22	9 (9)	3.00±0.21	0.256±0.056	0.336±0.051	0.262±0.101	0.239±0.064	0.274±0.067

Table 4.1. Multilocus descriptive statistics of head and body louse F_0 parental populations. Families, number of F0 parental pairs establishing F_1 broods. Loci, number of reliable loci included in analysis (informative i.e. polymorphic loci in parentheses). F_1/family, range of number of individuals per family genotyped for each ecotype. Allele/locus, mean number of alleles per marker ± SE. H_O, mean observed heterozygosity across all loci ± SE. H_E, mean expected heterozygosity across all loci ± SE. FIS, mean FIS across all loci ± SE (following Weir and Cockerham, 1984). H_O ♂ and H_O ♀, mean observed heterozygosity across all loci ± SE for F0 fathers and F0 mothers.

T9_6 was monomorphic in head lice, whereas T9_6 and T1_4 failed to amplify in most body louse individuals and were excluded in this ecotype, but all remaining loci were polymorphic and amplified consistently in most families. It was further decided that the T8_1 locus should be excluded in both ecotypes as a result of genotype inconsistencies. Therefore, from the initial 12-locus microsatellite panel, 10 and nine reliable informative loci were used to estimate allele transmission patterns in head and body louse families, respectively.

To assess the genome coverage of the microsatellite panel, all primer sequences were blasted to the *P. humanus* genome assembly to determine whether they were located in different genomic regions. Each of the markers was found to map to a distinct scaffold in the genome assembly (Table 4.2). Although the genome assembly does not allow for the exact determination of which chromosomes are targeted by the markers used herein, the present authors are confident that the panel offers sufficient coverage for a genomewide study of transmission patterns. By contrast, very limited genome coverage of the three markers used in McMeniman & Barker (2005) was found because two of them map to the same scaffold and the location of the third is unclear.

Table 4.2. Genome location of markers developed by Ascunce *et al.* (2013) (used in this study) and Leo *et al.* (2002) (used in McMeniman & Barker (2005)). All forward and reverse primers for each locus mapped to the same scaffold and the highest E-value for each of the pairs is shown, except for ML_10 (* F, ** R)

Panel	Locus	Scaffold	E-value
Ascunce *et al.*	M3_10	DS235157	0.002
	M3_19	DS235833	0.002
	M2_2	DS235048	0.0005
	T2_6	DS235090	0.0002
	M2_19	DS235875	0.0005
	M2_13	DS235785	0.002
	M2_3	DS235111	0.0005
	T9_6	DS235882	<0.0001
	T2_7	DS235100	0.002
	T4_5	DS235283	0.0002
	T1_4	DS235023	0.002
Leo *et al.*	ML_8	DS235886	<0.0001
	ML_9	DS235886	0.002
	ML_10	DS235042*	0.023
		DS235005**	0.98

4.4.2. Allele transmission patterns in males and females

For most families and loci in both ecotypes, heterozygous head and body louse males did not transmit alleles in a Mendelian fashion, but consistently passed on only one allele to the F_1. By contrast, females transmit both alleles to their offspring (Fig. 4.1; Appendix 4, Table S1). These patterns are consistent with PGE: females are normally diploid and exhibit Mendelian transmission, whereas males show whole-genome drive and transmit only the maternally inherited allele at each locus.

However, despite clear preferential transmission of one of the two alleles at each locus, head and body louse males sporadically also transmitted alternative (i.e. paternally inherited) alleles. Occasional paternal transmission of alternative alleles was observed across most markers, except for M2_19, M3_19 and M3_10 (Fig. 4.2).

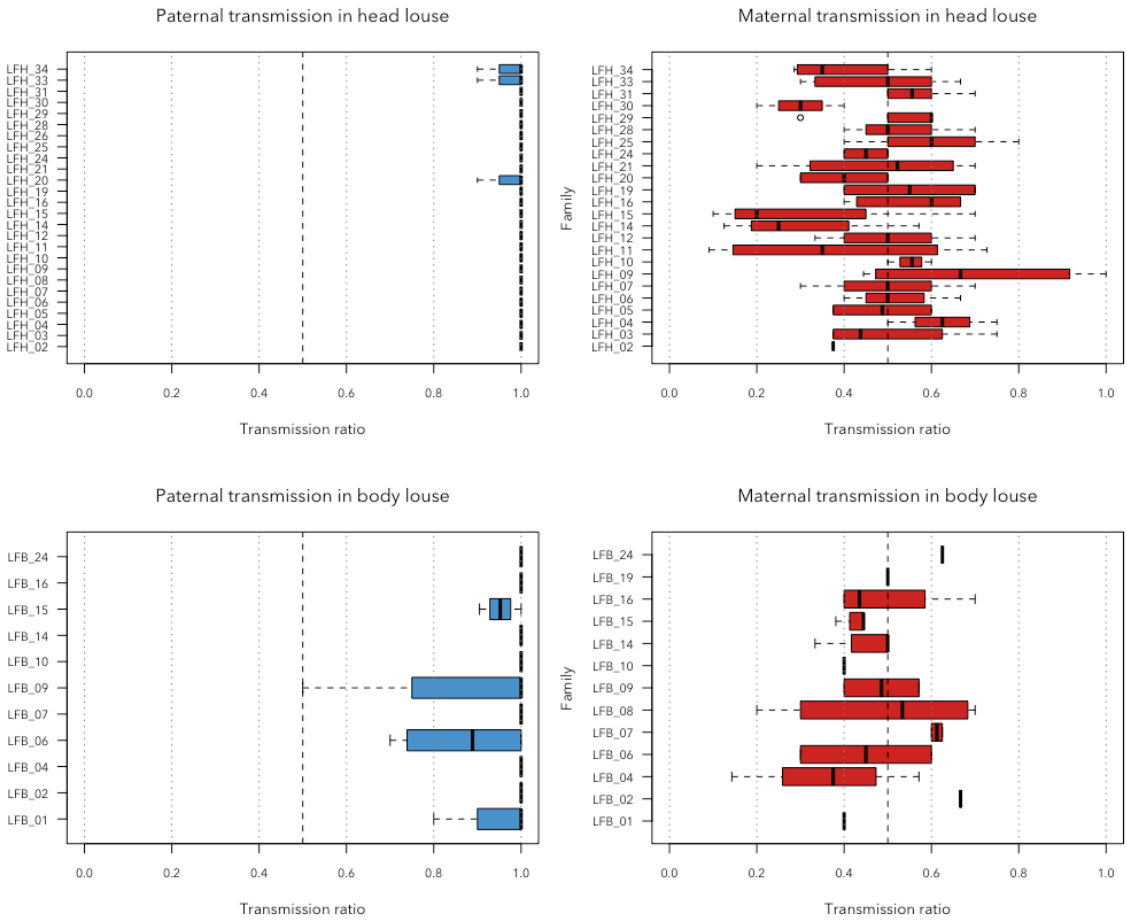


Figure 4.1. Allele transmission ratios across all loci for head and body louse males and females. When both alleles are equally represented in the F_1 offspring, transmission ratio is 0.5 (complete Mendelian transmission). A transmission ratio of 1 indicates complete drive of one of the parental alleles.

Escapes were rare in head louse males: three males (LFH_20, LFH_33 and LFH_34) passed on an alternative allele once at a single different locus (T1_4, M2_2 and T2_7, respectively). The other 23 head louse males showed complete PGE at all heterozygous loci. Overall, 64 of 71 head louse paternal transmission ratios deviated significantly from the Mendelian expectation of equal transmission at a significance threshold of 0.01 (all 71 at α=0.05), compared with one of 81 ratios in head louse females.

In body louse families, incomplete PGE occurrences were more frequent. Four of the 11 informative males also transmitted the alternative allele at least once (LBH_01 and LBH_09 at one locus, LBH_06 and LBH_15 at two loci). With a significance threshold of 0.01, 19 of 28 ratios deviated from Mendelian transmission (25 of 28 at α=0.05). In body louse females, none of the 29 transmission ratios deviated from Mendelian expectations.

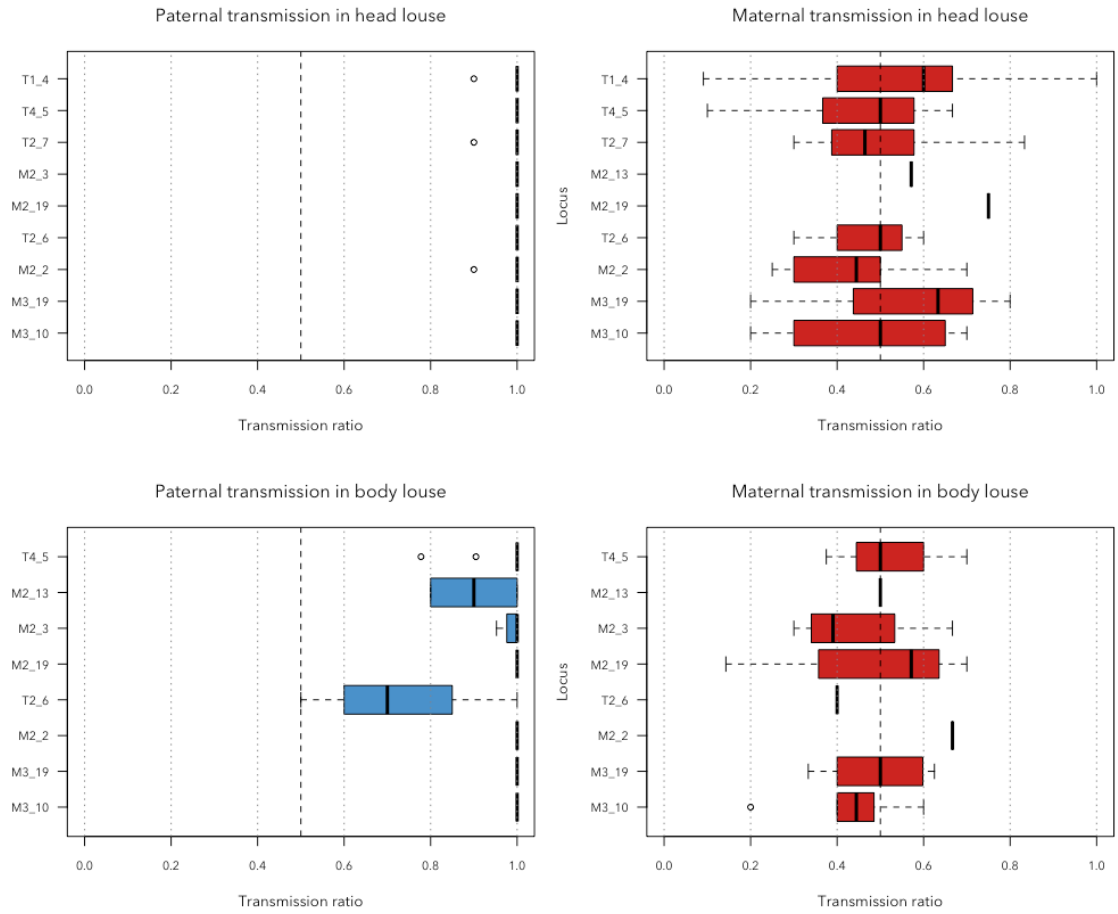


Figure 4.2. Paternal and maternal allele transmission ratios in all head and body louse families grouped by loci.

The present study did not find a consistent pattern of incomplete PGE instances across families and loci. To exclude genotyping error for these unexpected paternal escapes, both parents and offspring were re-genotyped and additional offspring were genotyped when available. We are therefore confident that the current findings represent true events of paternal chromosomes escaping germline elimination at low frequencies, particularly in body lice.

4.5. Discussion

The allele transmission patterns described in the present study offer conclusive evidence of a genome-wide male transmission ratio distortion in both ecotypes of *P. humanus*: males exclusively (or, in some cases, preferentially) transmit only one of their alleles to their offspring. In addition, heterozygous genotypes in males from both ecotypes unambiguously indicate that males are diploid and that both paternally and maternally inherited chromosomes are kept in the soma. Although the two-generation experimental design used in this study does not explicitly allow for determination of the parental origin of alleles in F_0 individuals, McMeniman & Barker (2005) already demonstrated that driving alleles were always maternally inherited in body louse males. All these findings are consistent with germline PGE, a pseudohaplodiploid genetic system in which males develop from fertilized eggs and are diploid, but eliminate chromosomes of paternal origin from their germline. This type of reproduction is also found in several other insect taxa such as mealybugs, the coffee borer beetle, two dipteran clades and booklice (Burt & Trivers 2006; Gardner & Ross 2014; de la Filia *et al.* 2015; Hodson *et al.* 2017).

All males in the present study exhibited whole-genome transmission ratio distortion with sporadic, inconsistent leakages of non-driving alleles in some individuals. Interestingly, the current data reveal that leakages are more frequent in body than in head lice, although the power to detect these occurrences was greater in the latter because twice as many head louse families were screened and they showed higher levels of heterozygosity. However, the study found no evidence of a female PGE-inducing genetic polymorphism as suggested by McMeniman & Barker (2005). In their

model, a codominant maternally transmitted genetic element is responsible for elimination of paternal alleles in male offspring so that females that are heterozygous for this element produce PGE sons that pass on only maternal alleles and non-PGE sons that transmit parental alleles in a Mendelian fashion. However, the mapping of markers to the louse genome revealed that McMeniman & Barker (2005) appear to have targeted a single chromosome only. Therefore, an alternative interpretation of these earlier results that is consistent with the sporadic leakage of paternal alleles observed in the current study would be a germline PGE mechanism in which discrimination between maternal and paternal chromosomes during spermatogenesis is not infallible. In germline PGE, males are somatically diploid and elimination of paternal chromosomes is achieved via non-random assortment of chromosomes during meiosis so that only nuclei containing maternal chromosomes develop into viable sperm (Burt & Trivers 2006). Whole-genome meiotic drive of maternal chromosomes in germline PGE taxa has been most extensively described in sciarid flies (Esteban *et al.* 1997; Goday & Esteban 2001) and mealybugs (Bongiorni *et al.* 2004; Bongiorni *et al.* 2009). Allele transmission patterns in louse males reveal that paternal chromosomes are similarly excluded from active spermatocytes, but are occasionally able to escape elimination by migrating with other maternal chromosomes in lieu of their homologues, particularly in body lice. Achiasmatic male meiosis, which is an imperative requisite for PGE as it prevents mixing of paternal and maternal alleles, has been documented in lice (Tombesi & Papeschi 1993; Tombesi *et al.* 1999; Bressa *et al.* 2015). As recombination between maternal and paternal homologues cannot account for transmission of paternal alleles, the detected leakage would encompass entire paternal chromosomes. Therefore, the apparent non-PGE body louse males found by McMeniman & Barker (2005) are more likely to be males exhibiting biparental transmission for the chromosome targeted by their marker panel only, whereas transmission of other chromosomes consistent with PGE (or additional occurrences of paternal leakages) would have passed undetected.

Head and most particularly body lice are the first species for which incomplete (albeit not polymorphic) PGE has been explicitly reported. The discrimination mechanism by which paternal and maternal louse chromosomes are differentially tagged is unknown. In other PGE taxa, maternal and paternal chromosomes differ in patterns of DNA methylation (Bongiorni *et al.* 1999; Bongiorni *et al.* 2009) and histone modifications

(Goday & Ruiz 2002; Greciano & Goday 2006; Khosla *et al.* 2006; Escribá *et al.* 2011; Prantera & Bongiorni 2012), which may mediate discrimination between homologues during spermatogenesis. In lice, inaccuracies of the parent-of-origin discrimination mechanism, whichever its nature, could result in the occasional migration of paternal chromosomes with the remaining maternal chromosomes.

Although at this stage the issue of how these leakages occur is subject only to speculation, a likely PGE mechanism in which only nuclei containing maternal chromosomes develop into viable sperm (bar accidental leakage of paternal homologues) can be proposed based on previous cytogenetic work in lice. Louse spermatogenesis is highly complex: achiasmatic meiosis is followed by three or four mitotic divisions to yield a 32/64-cell cyst that undergoes a final and unequal mitosis in which most cytoplasmic material is excluded from half the cells, which degenerate into pyknotic nuclei (Hindle & Pontecorvo 1942; Bressa *et al.* 2015) similar to those seen in mealybug spermatogenesis (Bongiorni *et al.* 2004; Bongiorni *et al.* 2009). The present authors agree with McMeniman & Barker (2005) that non-random assortment of chromosomes is likely to occur in the last, unequal division, after which only the spermatids carrying maternal chromosomes develop into viable spermatozoa. This implies an inverted meiotic sequence in which the first division is equational rather than reductional, with sister chromatids separating before segregation of homologous chromosomes, as found in other PGE taxa such as mealybugs (Viera *et al.* 2008). It is possible that inverted meiosis in louse males has been historically overlooked in cytogenetic studies as a result of the lack of heteromorphic bivalents and the tight association and highly condensed nature of louse chromosomes, which are holocentric [i.e. they lack a localized centromere; see Bressa *et al.* (2015) and references therein]. Recently, Bressa *et al.* (2015) reported that sister chromatid separation indeed occurs in the first division, but conclusive evidence has yet to be presented.

PGE may have important implications for the transmission of pesticide resistance, which must be parent-of-origin-dependent in males. Resistant males are unable to pass on the trait to their offspring when it is paternally derived and hence resistance will be lost through the paternal line even if it is under strong positive selection. By contrast, males that inherited the resistance trait from their mothers will transmit it to all their offspring, rather than half as occurs in Mendelian inheritance. These

characteristic PGE inheritance patterns complicate predictions of resistance invasion without models that explicitly consider sex-specific differences on allelic transmission. In addition, PGE also reduces effective population sizes (Wright 1933), although this effect may be small when sex ratios are female-biased (Hedrick & Parker 1997), as is often the case in louse populations (Perotti *et al.* 2004).

Another way in which PGE can affect the evolution of resistance is through its potential effect on patterns of gene expression. Taxa in which PGE occurs vary in the degree of paternal genome expression in males, which can affect response to insecticides and have an impact on rates of resistance evolution. In many PGE groups, paternal chromosomes are lost (haploid soma PGE) or transcriptionally inactive (functionally haploid PGE) (Normark, 2003) One immediate consequence of these two forms of PGE is that maternally inherited recessive alleles are directly exposed to selection in males, as under arrhenotoky. Therefore, the evolution of insecticide resistance is faster in arrhenotokous (Crozier 1985; Havron *et al.* 1987; Caprio & Hoy 1995; Denholm *et al.* 1998) and functionally haploid PGE species (Brun *et al.* 1995) than in diplodiploids [but not always; see (Carrière 2003)]. However, males in other PGE taxa are diploid and may express both alleles regardless of parental origin (diploid soma PGE) (Gardner & Ross 2014).

Because of this variation in gene expression patterns in PGE systems, it is important to precisely determine the degree of paternal genome expression in louse males. Although heterozygous males show that paternal chromosomes are retained, it is still possible that these are transcriptionally inert. In functionally haploid PGE taxa that remain somatically diploid, inactive paternal chromosomes appear as highly compacted dots (Brown 1972; Brun *et al.* 1995; Hodson *et al.* 2017). To the present authors' knowledge, this conspicuous chromosomal behaviour has never been described in human lice, which suggests that PGE is of the diploid soma form and paternal chromosomes are hence transcriptionally active. Phenotypic assays in hybrid individuals are other indicators of paternal chromosome expression in PGE males because they are expected to show the same traits as males from the maternal species if paternal chromosomes are inactivated. Body size and tibia length measurements in hybrids have been reported to be intermediate between head and body lice (Busvine 1978), but this study did not discriminate between male and female offspring.

If paternal chromosomes are expressed in human louse males, the aforementioned theoretical models on evolution resistance in haplodiploids cannot be applied because they do not consider diploid expression in PGE species with arrhenotokous-like inheritance. Therefore, new theory must be developed to predict how whole-genome meiotic drive in males with diploid gene expression will affect resistance evolution.

How PGE evolved in the human louse remains an open question. Although *P. humanus* is the only anopluran (i.e. sucking louse) in which the occurrence of PGE has been explicitly demonstrated, the same modified spermatogenesis has been reported in other parasitic louse species. These include another anopluran, the pig louse *Haematopinus suis* (Phthiraptera: Haematopinidae) (Bayreuther 1955; Tombesi & Papeschi 1993), and members of two suborders of the paraphyletic group Mallophaga (i.e. chewing lice): Amblycera [the guinea pig louse *Gyropus ovalis* (Phthiraptera: Gyropidae) and the chicken body louse *Menacanthus stramineus* (Phthiraptera: Menoponidae)] (Scholl 1955; Tombesi & Papeschi 1993) and Ischnocera [two species of *Bovicola* (Phthiraptera: Trichodectidae), the goat louse] (Tombesi & Papeschi 1993). More tellingly, empirical evidence of PGE in a close relative of parasitic lice, the booklouse *Liposcelis* sp. (Psocoptera: Liposcelididae), has been recently provided (Hodson *et al.* 2017) In this species, PGE is of the functionally haploid type, with males retaining condensed paternal chromosomes. Although the phylogenetic relationships between and within Psocoptera (booklice) and Phthiraptera are not yet fully resolved and this division has been called into question (Yoshizawa & Johnson 2010; Li *et al.* 2015), there is consensus that all lice form a monophyletic group and it is therefore possible that PGE is common to all of them. Formal investigation of transmission patterns and somatic heterochromatinization in these or other parasitic louse species would be necessary to corroborate this hypothesis.

Several authors have suggested that PGE may have evolved through attempts by endosymbionts to manipulate their host's reproduction (Normark 2004; Kuijper & Pen 2010; Ross *et al.* 2012). The rationale for this is that maternally transmitted endosymbionts benefit from a female-biased sex ratio and that the elimination of paternally derived chromosomes in males may be a way of killing male offspring. Lice harbour several maternally inherited endosymbiotic bacteria including both obligate

nutritional mutualists as well as bacteria known to manipulate host reproduction in their own favour, such as *Wolbachia*. Hence, could PGE in lice be induced by endosymbionts? Probably not: human louse males remain diploid throughout development and only eliminate their paternally derived genome during spermatogenesis, which is unlikely to induce male-specific mortality and is therefore not in the interest of the endosymbionts.

The present study demonstrates that PGE is present in both *P. humanus* ecotypes and outlines some considerations of the impact of the particular genetic system on the evolution of pediculicide resistance. A more complete understanding of human louse biology is imperative to facilitate the design and application of successful resistance management programmes. Yet asymmetry in gene transmission patterns, sex ratio bias and possible phenotypic consequences of PGE have not been considered thus far. The characterization and compact nature of the *P. humanus* genome enable genome-wide allele-specific expression studies to determine the extent to which paternally inherited alleles can confer resistance phenotypes in males. If they can, theoretical models of resistance evolution combining diploid expression and haplodiploid transmission will be needed to maximize the success of resistance control strategies.

Chapter 5

Parent-of-origin-specific transcriptome analysis in males of the human louse *Pediculus humanus*

5.1. Summary

Both ecotypes of the human louse *Pediculus humanus* (Phthiraptera: Pediculidae), the head and body lice, reproduce through paternal genome elimination (PGE), a form of haplodiploidy whereby males are diploid but transmit maternally-inherited chromosomes only, while the paternal homologues are excluded from sperm. Under PGE, these asymmetric inheritance patterns co-occur with somatic adaptations in the male soma to reduce or prevent expression of paternal alleles, such as transcriptional silencing or partial or complete embryonic elimination of paternal chromosomes. Predominant or completely haploid expression of maternally-inherited alleles has been linked to an evolutionary arms race between parental genomes and has profound consequences for males, including rates of response to selection. To date, no such effects have been reported in lice. In this study, a parent-of-origin allele-specific transcriptome analysis is used to establish to what extent paternal chromosomes are expressed in individual louse males deriving from crosses between head and body lice. Additionally, total gene expression levels between males from the parental ecotypes and between reciprocal hybrids are compared to determine if inter-ecotype males and pure species males show analogous differences, as expected under silencing of paternal alleles. Globally, patterns of allele-specific and total gene expression reveal that expression is biparental in this species, although a subset of genes show parent-of-origin effects that are consistent with genomic imprinting. These results are discussed in the context of PGE evolution, for which *P. humanus* could constitute a promising novel study system to understand the role of intragenomic conflict in early evolutionary stages of this bizarre genetic system.

5.1. Introduction

Paternal genome elimination (PGE) is a genetic system widely distributed across arthropods, in which males develop from fertilised eggs but their paternally-inherited

chromosomes are eliminated before or during spermatogenesis. Therefore, PGE males only transmit the maternal set to the offspring (Normark 2003; Burt & Trivers 2006; Gardner & Ross 2014). This transmission asymmetry is common to all 10,000 species where PGE is estimated to be present (de la Filia *et al.* 2015). Another striking feature of this genetic system is the variability in contribution of paternal chromosomes to genetic activity in the male soma (Herrick & Seger 1999; Gardner & Ross 2014). In PGE species where elimination of paternal chromosomes takes place early in development (embryonic PGE), such as mesostigmatid mites (Nelson-Rees *et al.* 1980; Norton *et al.* 1993) or diaspidid scale insects (Nur 1980; Ross *et al.* 2010a), males become true haploids and therefore only express maternal alleles. In some PGE species where elimination is delayed until spermatogenesis (germline PGE), such as Scolytinae beetles (Brun *et al.* 1995), neococcid scale insects (Brown & Nur 1964; Bongiorni & Prantera 2003; Ross *et al.* 2010a) or booklice (Hodson *et al.* 2017), paternal chromosomes are retained throughout development, but become tightly condensed during embryogenesis. As a result, paternal chromosomes remain transcriptionally inactive (Berlowitz 1965) and expression of maternal alleles predominates—but not always completely (Chapter 2). Other germline PGE taxa, such as symphypleonid springtails and the dipteran PGE families Sciaridae and Cecidomyiidae, lack this conspicuous behaviour of paternal chromosomes (Burt & Trivers 2006). However, in these species at least a fraction of the genome is haploid in males, as paternally-inherited X chromosomes are eliminated in early development (Goday & Esteban 2001; Benatti *et al.* 2010). One of the most intriguing aspects of the evolutionary history of PGE is the repeated convergence towards these adaptations to reduce or completely prevent expression of paternal chromosomes in males in the different groups where PGE has independently emerged.

To date, little is understood about the sources of partial or complete haploidisation of gene expression in males under PGE. The leading hypothesis assumes that whole-genome drive of maternal chromosomes evolved first (Brown 1964; Bull 1979; Haig 1993; Gardner & Ross 2014) and subsequent haploidisation of males represents a maternal genome adaptation to prevent resistance of paternal alleles to germline elimination (Herrick & Seger 1999; Ross *et al.* 2010a). Under this view, the evolution of silencing and embryonic elimination of paternal chromosomes represents different stages of a coevolutionary arms race between parental genomes over transmission to

the offspring. An empirical exploration of this hypothesis would benefit from a comparative approach to manifestations and mechanisms of PGE in a wider range of species representing different stages of this arms race.

One of such PGE species of special interest could be the human louse *Pediculus humanus.* In Chapter 4 (de la Filia *et al.* 2017), I showed that both ecotypes of this species, the head and body lice, reproduce via PGE, following an earlier report of inheritance patters consistent with this genetic system (McMeniman & Barker 2005). To current knowledge, PGE in human lice seems to be unique. Similarly to other PGE taxa, males have a strongly modified spermatogenesis: the first meiotic division is followed by a series of mitosis to form a 32/64 cell cyst, of which only half (those containing maternal chromosomes) continue to develop as active spermatozoa, while the other degenerate *in situ* (Hindle & Pontecorvo 1942; Bressa *et al.* 2015). However, none of the somatic phenomena associated to PGE—heterochromatization, elimination of single paternal chromosomes or the whole paternal set during development—have ever been described in human lice, despite extensive cytogenetic research. It is therefore logical to assume that expression in human louse males is fully diploid, although direct proof is lacking. From an evolutionary perspective, human lice could then represent the most basal form of PGE under the evolutionary arms race hypothesis, prior to the emergence of any putative maternal adaptations to counteract resistance of paternal alleles to germline elimination (Herrick & Seger 1999; Ross *et al.* 2010a), consistently with the high rate of paternal allele leakage found in the previous chapter. If this is the case, *P. humanus* would be a promising system to directly search for genes or chromosomal regions with parent-of-origin-specific expression which could be directly involved in the induction—or deterrence—of paternal chromosome elimination.

Moreover, somatic patterns of gene expression in *P. humanus* are also interesting beyond fundamental theories on intragenomic conflict and evolution of non-canonical genetic systems. From an applied perspective, a precise understanding of how PGE (or other asymmetric genetic systems) shapes key aspects of species biology, including sex-specific patterns of gene expression, is advantageous: for example, to help understand and predict selective responses in PGE taxa of economic and epidemiological relevance. Many PGE species with paternal chromosome silencing are widespread pests that pose a severe economic burden on crop production (Gill &

Kosztarab 1997; Damon 2000). PGE with haploid male expression has been shown to affect response to insecticides and accelerate resistance evolution, through direct exposure of maternally-inherited recessive alleles to selection (Brun *et al.* 1995). *P. humanus* is a widespread human ectoparasite with serious consequences (Clark *et al.* 2013): body lice constitute a serious health threat as vectors of severely pathogenic bacteria (Raoult & Roux 1999) and head lice infestations have been estimated to cause costs of hundreds of million dollars every year (Hansen & O'Haver 2004). In the last decades, increasing resistance to available pediculicides has become a major challenge to human louse control (Burgess 2004; Durand *et al.* 2012; Clark *et al.* 2015; Clark 2018), creating the need for novel treatments and resistance-proof management programmes. As I suggested in the previous chapter, the hitherto overlooked transmission patterns of PGE in both ecotypes could inform novel anti-pediculide strategies. Yet a full characterization of the form of PGE present in *P. humanus* is necessary to determine whether existing models of resistant evolution in haplodiploid taxa (Crozier 1985; Caprio & Hoy 1995; Denholm *et al.* 1998; Carrière 2003) can be applied.

In this chapter, I apply the same framework developed in Chapter 2 (Wang & Clark 2014) to analyse parent-of-origin allelic specific expression (ASE) patterns in F1 male offspring of crosses between head and body lice. Head and body lice are two distinct ecotypes of the same species, differing in morphological and behavioural traits driven by ecological factors (Light *et al.* 2008; Veracx & Raoult 2012). Consequently, the transcriptomic profiles of both ecotypes are highly similar, with low levels of inter-ecotype divergence in nucleotide sequences and gene expression levels (Olds *et al.* 2012). Therefore, individuals from both ecotypes can be easily crossed in laboratory conditions, yielding viable and fertile offspring (Busvine 1978). For this study, fixed discriminant SNPs between body and head louse ecotypes were called using DNA-seq and F1 transcriptomes were generated for single males derived from different mating pairs in reciprocal crosses between both ecotypes. To determine whether paternal alleles are expressed, ASE patterns were estimated for 162 genes, of which only a minority showed parent-of-origin-specific expression in reciprocal F1s. The results of this analysis confirm that human lice show biparental expression and open a promising avenue for the exploration of the dynamics of PGE.

5.2. Materials and methods

5.2.1 Experimental populations, inter-ecotype crosses and sample collection

Strains of head (HH) and body (BB) lice used in this experiment and rearing methods are as described in Chapter 4. Paternal generations (F0) were established by isolating 6 sexually immature third instar males and females from both colonies. F0 males from each ecotype were kept in common cages until sexual maturity, while F0 females were transferred to individual cages. Since sexually mature louse females can lay unfertised eggs that do not hatch (Bacot 1917), matings were delayed until females laid a first batch of eggs to confirm virginity. After 1-2 days, these eggs were removed from female cages to be incubated for a minimum of 10 days and a F0 male from the other ecotype was introduced. In total, 12 mating pairs were set: 6 HH ♀ x 6 BB ♂ and 6 BB ♀ x HH ♂. Females were allowed to lay eggs for 15 days before removal of the mating pair. None of the eggs laid by F0 mothers before mating hatched, so no crosses were discarded. In total, 4 HH ♀ x BB ♂ (HB1, HB2, HB3, HB6) and 4 BB ♀ x HH ♂ (BH1, BH4, BH5, BH6) were successful. Adult F1 males were collected in RNAlater after a 24h period of starvation.

In addition, to generate genomic and transcriptomic data from HH and BB ecotypes, adult individuals were directly isolated from the colonies (only males for RNA-seq, from both sexes for DNA-seq) and collected in RNAlater.

5.2.2. RNA and gDNA extraction and sequencing

RNA was extracted from a single F1 male per cross. Males were removed from RNAlater, washed twice in ice-cold 1X PBS and ground in 400 µl of TRIzol (Invitrogen). Total RNA samples were isolated with a PureLink RNA purification kit (Thermo Fisher Scientific, USA), purified with RNA Clean & Concentrator™-5 (Zymo Research, USA) and validated using the Bioanalyzer RNA 6000 Nano kit (Agilent). TruSeq stranded mRNA-seq libraries were generated by Edinburgh Genomics (UK) and sequenced on the Illumina NovaSeq platform (S2 flowcell, 50 bp paired-end reads).

In addition to F1 transcriptomes, I generated two replicates of RNA-seq data from 10 adult BB and HH males from the same strains. Both HH and one of the BB samples were sequenced on the Illumina HiSeq 4000 (75 bp paired-end). The second BB sample was sequenced with the F1 transcriptomes. For DNA-seq of parental strains, gDNA was extracted using a DNeasy Blood & Tissue kit (Qiagen, The Netherlands) from 10 adult individuals. TruSeq DNA Nano gel free libraries (350 bp insert) and sequencing on the Illumina HiSeq X (150 pb paired-end) were performed by Edinburgh Genomics.

5.2.3. Reference genome assembly and SNP calling

Reference genome and annotation used in this study correspond to the latest version of the body louse genome assembly JCVI_LOUSE_1.0 (Kirkness *et al.* 2010, downloaded from https://www.ncbi.nlm.nih.gov/genome/?term=pediculus+humanus on 02 Apr 2018). To identify SNPs private to parental ecotypes, I followed a modified version of the pipeline implemented in Chapter 2. After quality control and trimming, 112.6 million read pairs for BB and 127.9 million read pairs for HH were mapped against the JCVI_LOUSE_1.0 genome assembly using bwa 0.7.15-r1140 (Li & Durbin 2010). Raw variants were called using FreeBayes v1.1.0 (Garrison & Marth 2012) with the following settings: --haplotype-length 0 --standard-filters --min-alternate-count 2 --min-alternate-fraction 0.05 --min-coverage 10 --use-best-n-alleles 2 -J --pooled-discrete --pooled-continuous. As a first filter, non-SNP variants and SNPs supported by reads on single strands or mapping to one side of candidate variants were removed. A second filter was applied using bedtools v2.27.1 (Quinlan & Hall 2010) to remove variants detected in an ecotype that fell in regions without coverage in the other. To account for the expected reduced divergence between intraspecific parental genomes compared to hybrid crosses, I adopted a more relaxed criteria to call informative SNPs (Kincaid-Smith *et al.* 2018). SNPs with an alternative allele frequency of >90% in an ecotype and less than 10% of non-reference bases in the other ecotype were considered as fixed. Two distinct set of SNPs were generated after this final filter, each containing sites with alternate bases in BB or HH for which the opposite ecotype showed the JCVI_LOUSE_1.0 reference base.

5.2.4. RNA-seq mapping, maternal and paternal read counts and validation of ASE sites

Between 66.2-77.6 million RNA-seq read pairs were obtained for each F1 male sample. Read trimming and quality control were performed using fastp v0.18.0 (Chen *et al.* 2018) with the following settings: --cut_by_quality5 --cut_by_quality3 --cut_window_size 4 --cut_mean_quality 20 --trim_poly_g. On average, 74.4 million read pairs for each F1 sample were mapped to the JCVI_LOUSE_1.0 using STAR v2.5.2b (Dobin *et al.* 2013) in the two-pass mode. For pure BB and HH males, between 59.7-82.3 million RNA-seq reads were mapped to the reference genome. Normalisation and estimation of total gene expression levels for F0 and F1 males were performed using RSEM v1.2.28 (Li & Dewey 2011). Differentially expressed genes between pure F0 males and between reciprocal F1 males were detected using EBseq (Leng *et al.* 2013) with FDR<0.05.

To obtain allele-specific read counts at informative SNP positions, I used GATK v3.7 ASEReadCounter (McKenna *et al.* 2010) on all BH and HB F1 replicates and the merged BB and HH F0 transcriptomes. For all samples, allele-specific read counts were obtained separately for both BB and HH-specific SNP sets. Due to the paucity of informative SNPs and the additional power offered by a reciprocal cross design, the minimum read depth at each SNP position was reduced from 30 (Chapter 2) to 10. After obtaining allele counts, both sets of SNP positions were merged and it was arbitrarily decided to estimate allelic bias p at each SNP as the proportion of body louse allele reads (p = body louse reads/read depth at site). Therefore, in BH males, p_{BH} represents the maternal ratio in the F1 offsping of BB ♀ x HH ♂ and p_{HB} the paternal ratio in males of the reciprocal cross HB (Wang & Clark 2014).

Only sites present in all BH and HB replicates were kept after dropping SNPs in which at least a single sample showed >10% reads with bases other than reference and alternate. Then, BB and HH transcriptomes were used to validate SNPs by obtaining allele-specific counts in pure F0 males. Sites with p<0.9 in the pure BB male transcriptome and p>0.1 in HH were removed. Upon manual inspection of allele specific biases across F1 samples, I implemented a last filter to remove SNPs that showed strong inconsistencies between replicates of the same reciprocal cross. A G-test of

independence (with Bonferroni correction) was performed for each SNP position to test whether p_B values were homogenous across all 4 BH and HB replicates, and sites with significant heterogeneity in either or both reciprocal cross replicates were filtered out.

5.2.5. Allele-specific expression analysis

To estimate ASE patterns at SNP level, counts from BH and HB replicates were pooled. Following Wang & Clark (2014), SNPs with $0.35 \geq p_B \geq 0.65$ were classified as biparentally expressed (B). The remaining sites were considered putatively imprinted. In BH, SNPs with $p_{BH}>0.65$ were considered to show maternally biased (MB) expression (exclusively maternal (M) when $p_{BH}>0.95$) and paternally-biased (PB) when $p_{BH}<0.35$ (exclusively paternal (P) when $p_{BH}<0.05$). In HB, candidate imprinted sites were classified as MB with $p_{HB}<0.35$ (M when $p_{HB}<0.05$) and PB with $p_{HB}>0.65$ (P when $p_{HB}>0.95$).

SNPs were assigned to annotation features using asa2. For each gene with allele-specific information (at least a single SNP, with no depth requirement), p_B was estimated by pooling counts from all exonic SNPs in each replicate. A G-test of independence was performed across BH and HB replicates to test for homogeneity across reciprocal cross replicates. To detect parent-of-origin specific expression, Fisher's exact tests were performed for each gene by pooling parental counts across all BH and HB replicates (Wang & Clark 2014). Genes with significant allelic expression bias difference were assigned to a parental bias category using the same p thresholds as at SNP level, but incorporating reciprocal information: M when $p_{BH}>0.95$ and $p_{HB}<0.05$, MB when $p_{BH}=0.65$-0.95 and $p_{HB}=0.05$-0.35, PB when $p_{BH}= 0.05$-0.35 and $p_{HB}=0.65$-0.95 and P when $p_{BH}<0.05$ and $p_{HB}>0.95$. For genes for which p_{BH} and p_{HB} were not significantly different, *cis*-acting regulatory variant effects favouring expression of one of the alleles regardless of parental origin (*cis*-eQTLs consistent with lineage-of-origin effects) were considered when both p_{BH} and $p_{HB}>0.65$ (favouring the body louse allele) or p_{BH} and $p_{HB}<0.35$ (favouring the head louse allele). All other genes that did not show parent- or lineage-of-origin effects were considered to exhibit biparental expression. Information on putative functions and insect orthologs of genes showing parent-of-origin expression was gathered from VectorBase (https://www.vectorbase.org).

5.3. Results

To determine whether both parental genomes are transcriptionally active in *P. humanus*, and to which extent allele expression patterns conform to biparental expectations, parent-of-origin-specific ASE was estimated in whole adult F1 males originated from reciprocal crosses between the body (BB) and head (HH) louse ecotypes (Fig. 5.1). Transcriptomes were obtained for eight F1 samples, all of which originated from a different mating pair: 4 males produced in BB ♀ x HH ♂ crosses (BH) and 4 males produced in HH ♀ x BB ♂ crosses (HB). After trimming and quality control, between 66.2 and 80.3 million RNA-seq read pairs were mapped to the JCVI_LOUSE_1.0 reference genome to estimate allele-specific counts. 76%-80% of reads in each sample mapped uniquely to the reference genome, with duplication rates ranging between 52%-63%.

5.3.1. Informative ASE sites

To assign parental origin of RNA-seq reads, informative SNPs between BB and HH ecotypes were obtained by calling variants in both ecotypes separately against the reference genome. Due to the compact nature of the louse genome (110Mb), alignment of DNA-seq data from both ecotypes to the reference genome yielded very high coverage (230X in BB, 262x in HH), with >95% mapping rates. After removing low quality sites, 1,446,099 SNPs between BB and JCVI_LOUSE_1.0 and 1,172,339 between HH and JCVI_LOUSE_1.0 were called. Of these, 294,585 (20.4%) and 413,992 (35.3%), respectively, were considered fixed within each ecotype (proportion of alternative allele > 90%). The final set of informative SNPs was obtained by removing positions with no coverage or a reference allele frequency inferior to 90% in the other ecotype. In total, 92,520 SNPs (36,869 from BB and 55,651 from HH) were used to obtain allele-specific counts from F1 transcriptomes.

Between 6,865 and 8,837 of these SNPs were found in each of the 8 F1 male transcriptomes with a read depth of at least 10X. Of these, 5,727 SNPs were present in all BH and HB replicates, 59 of which (1%) were discarded for showing more than 10% of non-body and non-head louse bases in at least one of the samples (either due to

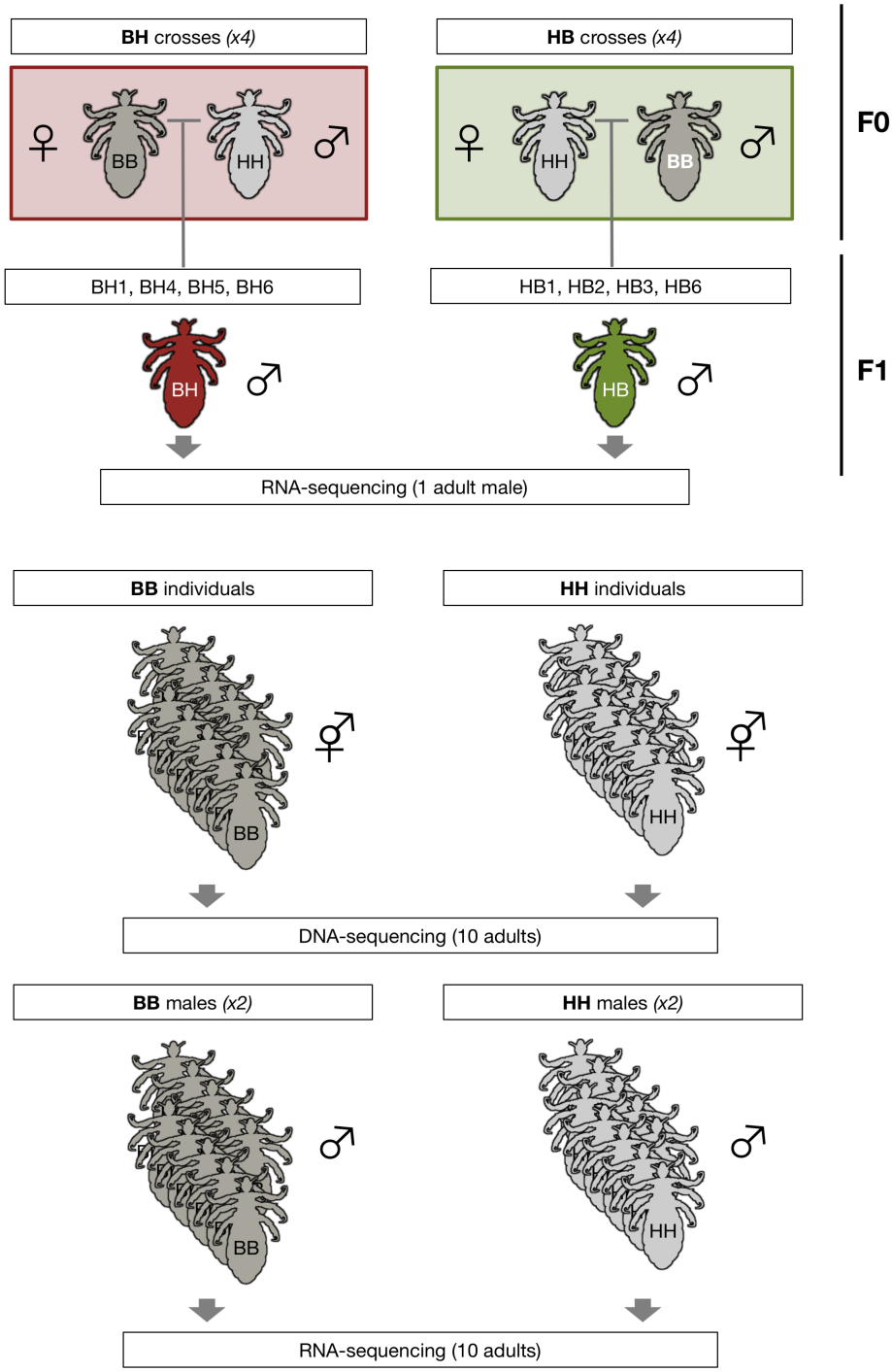


Figure 5.1. Experimental design. Above, reciprocal cross scheme between body (BB) and head (HH) louse ecotypes. F0 individuals were isolated from ecotype colonies and mated to individuals of the opposite sex from the other ecotype. 4 mating pairs from each reciprocal cross direction (BB ♀ x HH ♂ and HH ♀ x BB ♂) were successful. Transcriptomes were obtained from a single adult F1 male from each mating pair. For variant calling (below), a pool of 10 individuals from each ecotype was sequenced to obtain genomic reads. Additionally, to improve filtering of ASE sites, transcriptomes from two pools of 10 adult males from each ecotype were sequenced.

having a different allele in one of parental genomes or nucleotide mosaicisms). In parallel, allele counts for the same initial panel of 92,520 SNPs were obtained for merged transcriptomes from the two pools of pure BB (9,961 SNPs) and HH males (11,103). In BB, only 3,018 SNPs (30.1%) were found to be fixed (proportion of BB allele counts > 90%), in contrast to 7,398 (66.7%) in HH (proportion of HH allele counts > 90%). Non-fixed sites were removed from the F1 panel, which dropped to 936 SNPs. 1/3 of these sites were further removed as a result of significant heterogeneity across either BH or HB replicates (G-test, Bonferroni corrected p-value=5.3×10^{-5}). The final panel used to estimate ASE patterns at SNP and gene level consisted on 605 informative SNPs, of which 450 (74.4%) mapped to exonic regions in the JCVI_LOUSE_1.0 annotation (Appendix 5, Tables S1-S2).

5.3.2. Patterns of ASE at SNP and gene levels

For every informative ASE site in F1 males, BB and HH base counts were obtained to estimate parent-of-origin ASE. Allelic expression bias, p, was estimated as the proportion of BB reads at each informative SNP. In all samples, allelic expression bias conformed to biparental expression, with most SNPs showing values of p between 0.35-0.65 (Fig. 5.2A). In both BH and HB replicates, p was found to be moderately to strongly correlated between sample pairs, with higher correlations for read depth (Fig. 5.2B). For all replicates combined, no correlation was found between p_{BH} and p_{HB} at SNP level (Spearman's rank correlation coefficient ρ=0.004), with an absolute difference in allelic expression bias of 18% between reciprocal crosses (mean $|p_{BH}-p_{HB}|$=0.183). However, the correlation in combined read depth between reciprocal F1s was remarkably high (ρ=0.95). Most SNPs (60.3% in BH and 67.0% in HB) showed biparental expression in both reciprocal F1s (Fig. 5.3).

Many sites, however, showed allelic expression bias. 33.4% in BH and 19.3% in HB, respectively, were biased towards the maternal genome, with an additional 3.6% and 2.6% being completely expressed from the maternal genome. Only 1.8% of SNPs showed a bias towards the paternal genome in BH (and 0.8% were completely expressed from the paternal genome), while in HB 9.9% were biased to and 1.2% were exclusively expressed from the paternal genome.

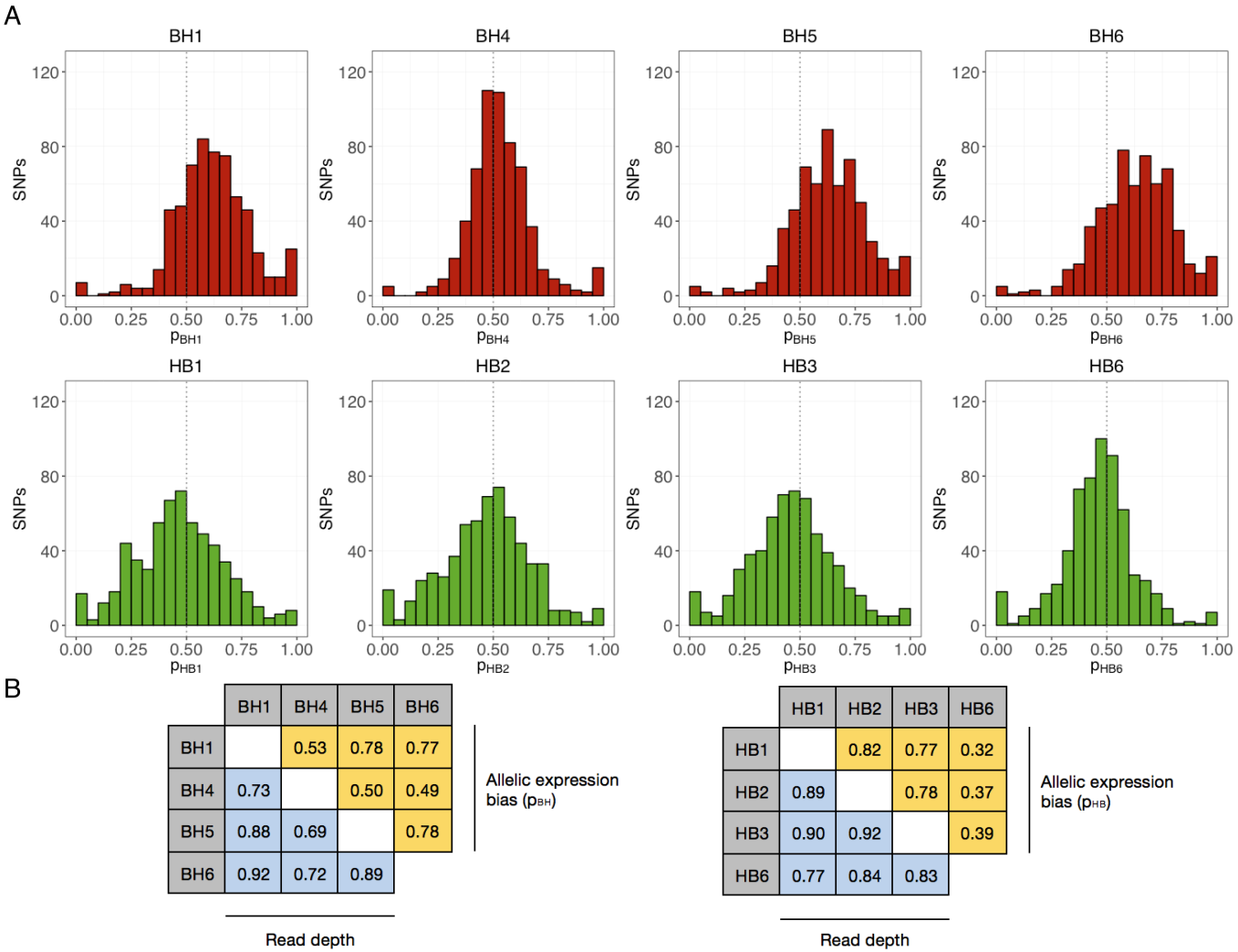


Figure 5.2. Allelic expression bias (p) in F1 BH (red) and HB (green) replicates. **(A)** Histogram of SNP counts. **(B)** Read depth and allelic expression bias correlation matrices between BH and HB replicates (Spearman's rho)

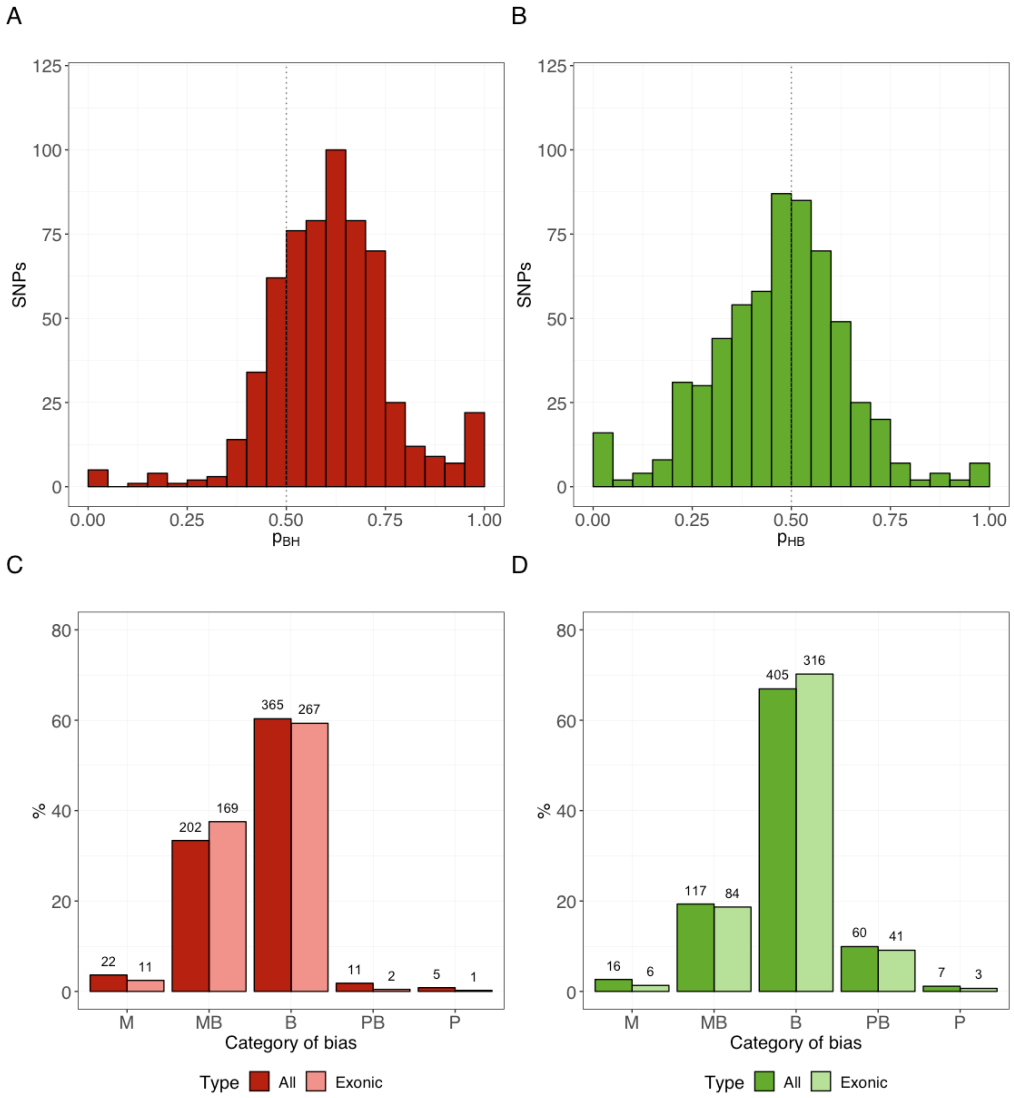


Figure 5.3. Combined SNP allelic expression bias (p_B). **(A, B)** Histograms of SNP counts in BH and HB, respectively. **(C, D)** Histograms of parent-of-origin specific bias category. M, exclusively maternal SNPs (p_{BH}>0.95, p_{HB}<0.05). MB, maternally biased (p_{BH}=0.65-0.95, p_{HB} =0.05-0.35). B, biparental (p=0.35-0.65). PB, paternally biased (p_{BH}=0.05-0.35, p_{HB}=0.65-0.95). P, exclusively paternal (p_{BH}<0.05, p_{HB}>0.95). Total SNP counts are shown in dark colour, exonic SNP counts in light colour.

Exonic SNPs were assigned to 162 genes with detectable expression levels (TPM > 0). On average, gene SNP density was 2.8 (SD 3.5), with 50% of genes being covered by at least 2 SNPs. For 126 of these genes (78%), both p_{BH} and p_{HB} were homogeneous across reciprocal cross replicates (G-test, Bonferroni corrected p-value=$3.1x10^{-4}$). After pooling counts from all 4 reciprocal cross replicates, correlation between p_{BH} and p_{HB} increased moderately at gene level (ρ=0.13), with an average reciprocal difference of 11%, and remained very high for read depth (ρ=0.98).

In total, 93 genes (57%) did not exhibit parent-of-origin-expression effects (Table 5.1, Fig. 5.4A), as p_{BH} and p_{HB} were not significantly different (Fisher's exact test, Bonferroni corrected p-value=3.1×10^{-4}). 80 of these genes showed complete biparental expression (average p_{BH}=0.56 and p_{HB}=0.52). The remaining 12 genes displayed

Table 5.1. Summary information and ASE patterns of genes with parent-of-origin specific information in lice. QTL, *cis*-eQTLs indicating lineage-of-origin effects

Genes with allele-specific information								
Total			*Homogeneous across replicates*			*Heterogeneous across replicates*		
162			126			36 (in BH, 19; HB, 9; both, 8)		
Allelic expression bias difference ($p_{BH} - p_{HB}$)								
Not significant								
Total			*Homogeneous across replicates*			*Heterogeneous across replicates*		
93			81			12		
B	QTL (BH)	QTL (HB)	B	QTL (BH)	QTL (HB)	B	QTL (BH)	QTL (HB)
80	12	1	74	7	0	6	5	1
Significant, consistent bias P_{BH} and P_{HB}								
Total			*Homogeneous across replicates*			*Heterogeneous across replicates*		
38			26			12		
B	MB	PB	B	MB	PB	B	MB	PB
26	10 (3)	2 (1)	18	6 (2)	2 (1)	8	4 (1)	0
Significant, different bias P_{BH} and P_{HB}								
Total			*Homogeneous across replicates*			*Heterogeneous across replicates*		
31			19			12		
B_{BH} and MB_{HB}	B_{BH} and PB_{HB}	B_{HB} and MB_{BH}	B_{BH} and MB_{HB}	B_{BH} and PB_{HB}	B_{HB} and MB_{BH}	B_{BH} and MB_{HB}	B_{BH} and PB_{HB}	B_{HB} and MB_{BH}
5	1	25	3	1	15	2	0	10

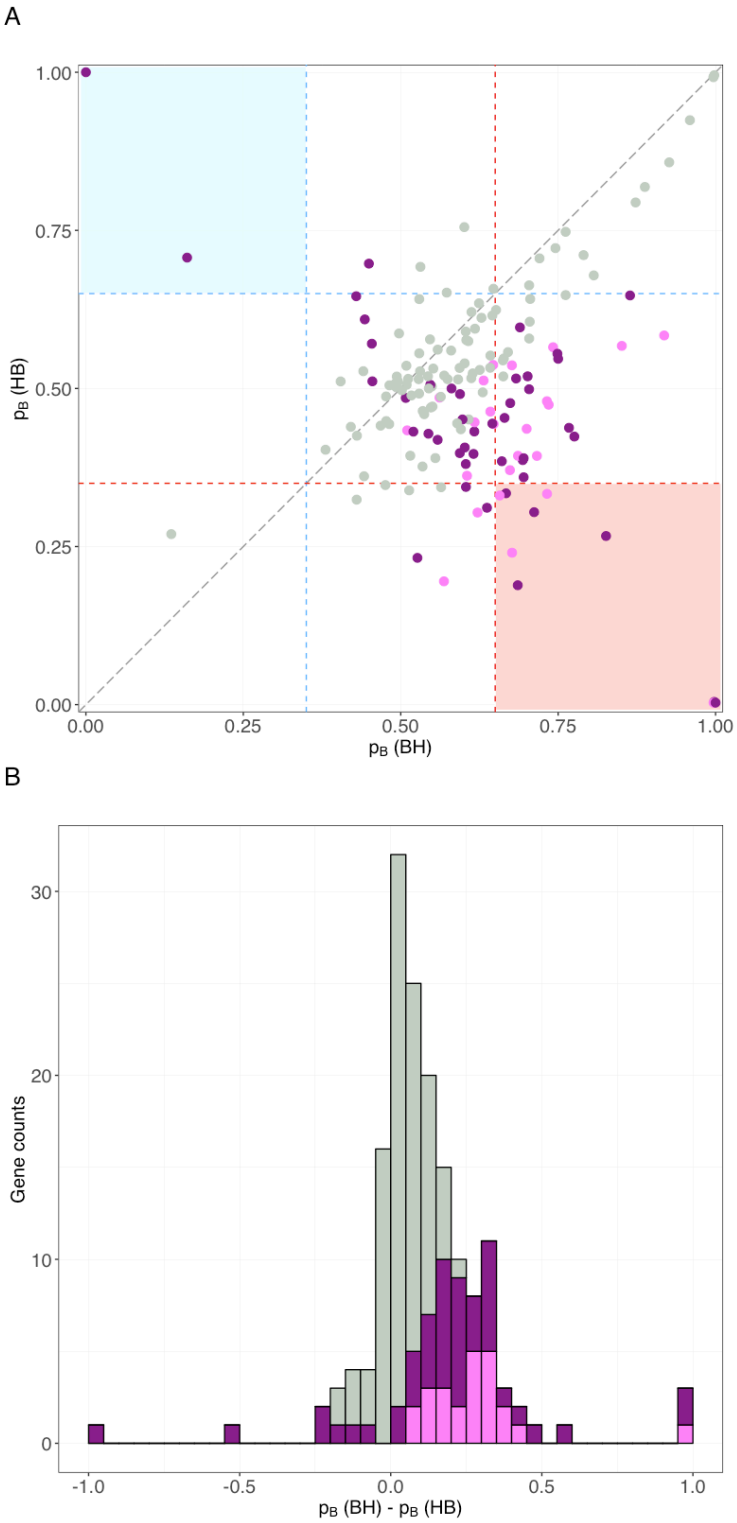


Figure 5.4. Combined allelic expression bias at gene level. Genes with significant allelic expression differences between BH and HB are shown in purple (when homogeneous across BH and HB samples) and pink (when not homogeneous across either BH or HB samples, or both) **(A)** Scatterplot of joint distribution of allelic expression (p) in F1 males. Genes with parent-of-origin-specific expression appear in the top left (paternally biased, shaded in blue) and bottom right areas (maternally biased, shaded in red). **(B)** Histogram of allelic expression differences ($p_{BH} - p_{HB}$)

lineage-of-origin-specific ASE patterns (*cis*-eQTLs) favouring expression of body louse (12 genes) or head louse genotypes (1 gene). Among the 69 genes with significant difference in reciprocal allelic expression bias (45 of which showed homogeneity between replicates), 26 were considered biparentally expressed (average p_{BH}=0.61 and p_{HB}=0.50), as they did not meet the criteria for differential parental expression (p_{BH} and p_{HB}=0.35-0.65). Another subset of genes with differential reciprocal allelic bias included 31 genes which displayed a parental bias in one of the F1 genotypes only, most of which (25) had a biparental pattern in HB males but showed a maternal bias in BH. Finally, the remaining 12 genes displayed complete parent-of-origin-specific expression patterns, with biased or exclusively maternal or paternal expression in reciprocal F1s. Overall (Fig. 5.4B), these gene expression patterns clearly indicate that both BH and HB F1 males have biparental expression—moderately biased towards the maternal genome—with a fraction of imprinted genes with parent-of-origin-dependent expression.

5.3.3. Location and function of genes with parent-of-origin-specific expression

Among the 12 genes with parent-of-origin-specific expression patterns, 7 have an assigned putative function in the JCVI_LOUSE_1.0 genome assembly (Table 5.2). Among the three completely maternal genes identified in F1 males, one is involved in fatty acid synthesis (acetyl-CoA carboxylation) and two have unknown functions. Among the maternally biased genes, 3 can be pinpointed for their putative roles in reproduction and differential allelic expression. The first of these genes in PHUM514040, a member of the MCM proteins family which act as key regulators of DNA replication licencing and initiation (Forsburg 2004). In *D. melanogaster*, its ortholog *mcm5* is additionally involved in meiotic recombination and chromosome segregation (Lake *et al.* 2007). The second gene, PHUM328760, is annotated as a Rad50-like double-strand break repair protein, a conserved member of a protein complex involved, among other roles, in recombination (Connelly & Leach 2002). However, the *D. melanogaster Rad50* 1-1 ortholog is a different gene, PHUM549380, which is also found among the genes with ASE information identified in this study (showing differential allelic expression bias but not a parent-of-origin expression pattern, with p_{BH}=0.66 and p_{HB}=0.45). PHUM328760's 1-1 ortholog in *D. melanogaster* is *Su(z)2*, a member of the Polycomb family involved in epigenetic regulation of gene expression (Kassis *et al.* 2017; Dasari *et al.* 2018).

Tellingly, *Su(z)2* is also maternally-expressed in *Drosophila*, at least in embryonic stages (Adler *et al.* 1989). The third maternally biased gene of putative interest, PHUM544500, is a cleavage and polyadenylation specificity factor whose *D. melanogaster* ortholog, CG7185, is involved in regulation of alternative mRNA splicing (Park *et al.* 2004). Finally, only one of the two paternally imprinted genes found in this study, PHUM395570, has an identified function and a 1-1 ortholog in *D. melanogaster.* This gene is a pro-resilin precursor, a highly elastic cuticle structural protein (Andersen 2010).

Next, the location of parentally biased genes was investigated to identify possible genomic regions displaying parent-of-origin expression. Among the 43 contigs to which genes with informative SNPs were mapped, 39 included at least one biparentally-

Table 5.2. Genes showing parent-of-origin specific expression. Bias, category of expression bias (M and P, completely maternal or parental; MB and PB, maternally- or paternally-biased). SNP, number of ASE sites within coding regions. p_{BH} and p_{HB}, allelic expression bias in BH and HB. Hom, homogeneous across samples.

Gene	Bias	SNP	p_{BH}	p_{HB}	Hom	Function	1-1 ortholog in *D. mel*
Maternally biased							
PHUM 622880	M	2	1.00	0.00	Yes	Acetyl-CoA carboxylase	-
PHUM 545480	M	3	1.00	0.00	No	Conserved hypothetical protein	-
PHUM 545470	M	1	1.00	0.00	Yes	Conserved hypothetical protein	-
PHUM 433690	MB	1	0.83	0.27	Yes	Purine permease	CG6293
PHUM 514040	MB	1	0.69	0.19	Yes	DNA replication licensing factor mcm5	*mcm5*
PHUM 060680	MB	1	0.68	0.24	No	Conserved hypothetical protein	CG4041
PHUM 328760	MB	1	0.71	0.30	Yes	DNA double-strand break repair Rad50 ATPase	*Su(z)2*
PHUM 545490	MB	27	0.73	0.33	No	Conserved hypothetical protein	-
PHUM 549400	MB	2	0.67	0.33	Yes	Pterin-4-alpha-carbinolamine dehydratase	*Pcd*
PHUM 544500	MB	10	0.66	0.33	No	Cleavage and polyadenylation specificity factor	CG7185
Paternally biased							
PHUM 395570	P	1	0.00	1.00	Yes	Pro-resilin precursor	*resilin*
PHUM 596750	PB	1	0.16	0.71	Yes	Abnormal long morphology protein	-

expressed gene, and 10 parent-of-origin genes were located on contigs where at least an additional gene displayed biparental expression (Appendix 5, Fig. S1, Table S3). Thus, these contigs housing imprinted genes did not display a parent-of-origin-specific expression pattern along their whole length. The remaining two parentally biased genes, PHUM622880 (M) and PHUM395570 (P), were the only genes with ASE information in their contigs (NW_002988572 and NW_002987464, respectively). Additionally, it was found that three adjacent genes with parent-of-origin-specific expression patterns are nested within introns of other known gene. PHUM545470 (M), PHUM545480 (M) and PHUM545490 (MB) are located within a 7.9 Kb intron of the PHUM545460 gene, a putative sodium-dependent nutrient amino acid transporter with a biparental expression pattern (p_{BH}=0.56 and p_{HB}=0.49).

5.3.4. Differential gene expression between pure males and F1 males

Differences in total gene expression were estimated for pure body and head louse males and between F1 males. In pure males, the correlation in gene expression between BB and HH was extremely high (ρ=0.99). Among 10,521 genes with detectable expression levels in at least one sample, only 550 (5.4%) were differentially expressed between head and body louse males (FDR<0.05) (Fig. 5.5A, Table S4 in Appendix 5). Of these genes, 252 were more highly expressed in body louse males and 298 were overexpressed in head lice. Between F1 males, expression levels were equally highly correlated between BH and HB (ρ=0.99) and the number of differentially expressed genes between genotypes was reduced to 74 out of 10,573 (0.7%) (Fig. 5.5B, Table S5 in Appendix 5). When comparing differential gene expression between pure males and between reciprocal F1s, F1 males did not reproduce the patterns in parental ecotypes. Fold changes in gene expression between BB/HH and BH/HB were found to be very weakly correlated (ρ=0.12) (Fig. 5.5C). Furthermore, very little overlap was found between sets of differentially expressed genes between pure males and between reciprocal F1s, as most overexpressed genes in BH and HB were not differentially expressed between body and head lice (Fig. 5.5D). Along this line, correlation coefficients in expression levels between pairs of pure and F1 males were consistently very high (ρ>0.90 in all pairs) (Fig. 5.5E).

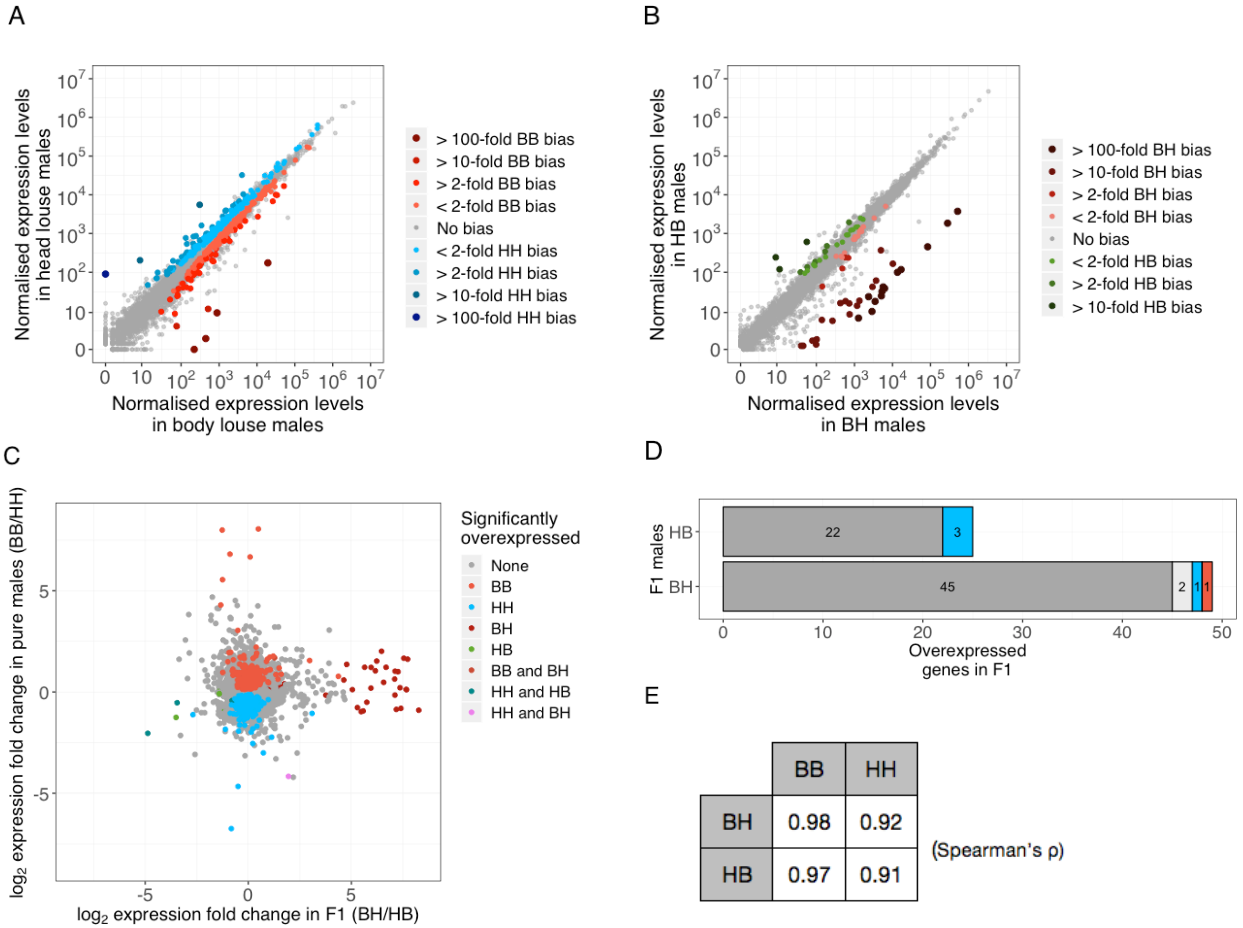


Figure 5.5. Total gene expression levels. **(A, B)** Scatterplot of normalised gene expression levels in pure BB and HH males (A) and BH and HB F1 males (B). Differentially expressed genes are coloured according to the fold differences in gene expression. **(C)** Scatterplot of fold expression changes between pure and F1 males. **(D)** Counts of overexpressed genes in F1 males and their overlap with differential expression genes between pure body (red) and head louse (blue) males (grey, not differentially expressed between parental ecotypes) **(E)** Correlation matrix in total gene expression levels between pure and F1 genotypes.

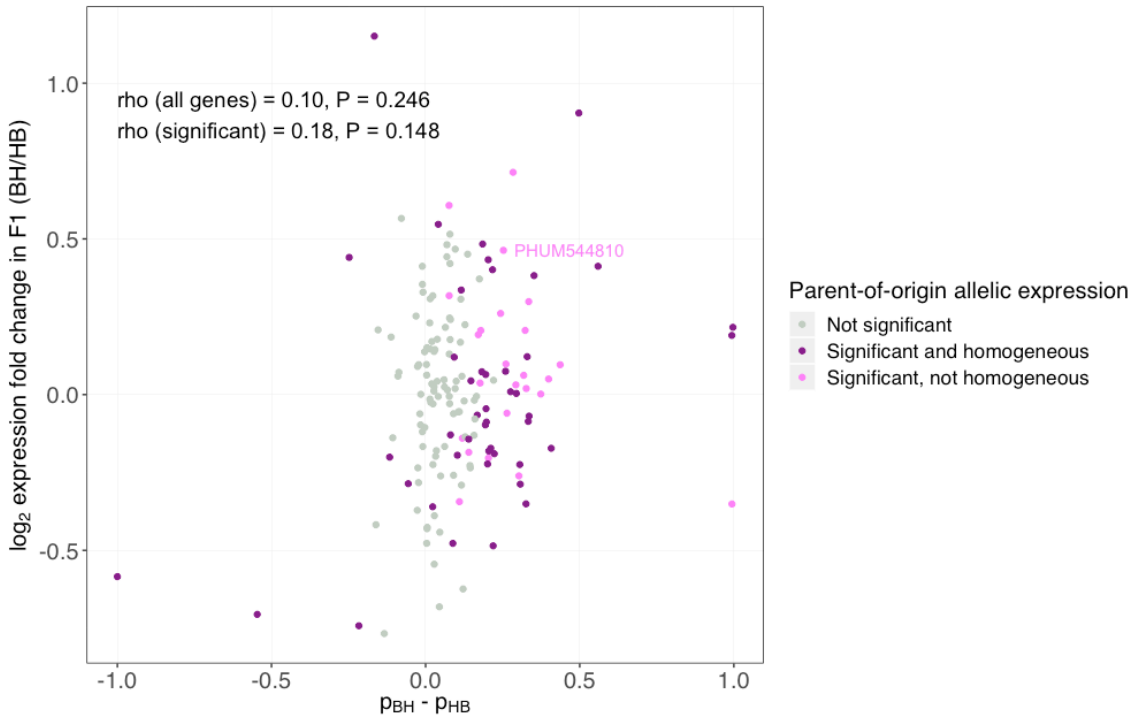


Figure 5.6. Relationship between fold changes in expression levels between F1 males and reciprocal allelic expression differences. Genes with significant parent-of-origin expression are coloured as in Fig. 5.3.

Finally, the relationship between ASE patterns and gene expression in F1 males was evaluated (Fig 5.6). Among differentially expressed genes between BH and HB, only one had an informative SNPs allowing quantification of ASE (PHUM544810). All the remaining 151 genes for which ASE was estimated showed equal expression patterns between BH and HB. Globally, no parent-of-origin effects in gene expression differences between F1 males were detected, as reciprocal allele expression differences (p_{BH} - p_{HB}) and gene expression changes between BH and HB males were not correlated (ρ=0.10 for all genes and rho=0.18 for genes with significant p_{BH} - p_{HB} only; P>0.05 in both cases).

5.4. Discussion

After describing PGE transmission patterns in head and body lice in the previous chapter, a RNA-sequencing based allele-specific expression study was a logical follow-up to fully characterise the form of PGE present in this species. In the past decade, RNA-

sequencing based allele-specific expression analyses have become increasingly widespread beyond their initial implementation in model organisms (Babak *et al.* 2008; Wang *et al.* 2008; Heap *et al.* 2009; McManus *et al.* 2010; Gehring *et al.* 2011). These approaches have contributed to our understanding of different aspects of genome function and gene regulation across a variety of taxa. In particular, ASE studies are a powerful tool to quantify and decompose *cis*- and *trans*-acting components of gene regulatory divergence between species (Wittkopp & Kalay 2012; Bell *et al.* 2013), describe patterns of X chromosome inactivation and dosage compensation (Lott *et al.* 2011; Wang *et al.* 2012; Marks *et al.* 2015) and identify genomic imprinting (Wang & Clark 2014). In this chapter, I investigated patterns of parent-of-origin-specific expression in transcriptomes of reciprocal F1 males to determine whether maternally-inherited alleles are predominant—which, to current knowledge, is the case in all other PGE species to different extents—or if expression is biparental genome-wide, as suggested by the lack of cytogenetic findings suggesting otherwise. Although hampered by the limited number of fixed discriminant SNPs between the parental genotypes, these results clearly show that both parental chromosomes are expressed in louse males, but identified a subset of maternally-biased genes which could be candidates for downstream analysis to investigate their possible involvement in exclusion of paternal chromosomes from the germline.

5.4.1. Complete biparental parental expression is predominant in louse males

According to the estimations obtained in this study, around two thirds of informative SNPs are expressed in a non-allelic specific manner in louse males, in line with other estimates of biallelic SNP expression in invertebrate diploid taxa (Kincaid-Smith *et al.* 2018). Consistently with SNP patterns, 65% of informative ASE genes exhibited complete biparental expression in both BH and HB males. A further 8% of genes showed complete or predominant monoallelic expression depending on lineage-of-origin, rather than parent-of-origin. These asymmetric expression patterns are probably caused by *cis*-acting regulatory variants favouring the expression of the same allele (mostly that of BB-origin) in both reciprocal crosses (Pollard *et al.* 2008; Wang & Clark 2014). Together, expression of ~3/4 of genes to which informative SNPs could be assigned does not exhibit parent-of-origin effects. These results are in stark contrast to

the ASE patterns obtained in Chapter 3 for hybrid mealybug males (where paternal chromosomes are heterochromatised), in which less than 1% of genes were found to be biparentally expressed.

Additional evidence for transcriptional activity of both maternal and paternal alleles in lice is provided by comparisons of differences in total gene expression levels between genotypes. Under PGE with suppression of paternal allele expression, F1 males are expected to be phenotypically identical to males from the maternal line, even under incomplete paternal genome silencing (see for example Brown & Nur 1964; Brown & Wiegmann 1969; Brun *et al.* 1995; Borsa & Kjellberg 1996). Therefore, if paternal chromosomes were not expressed in louse males, reciprocal F1s should display the same differences in gene expression as pure head and body louse males. In this study, I identified >500 differentially expressed genes between F0 males. Previous evaluations of differentially expressed genes between pools of head and body lice from all developmental stages (instead of adult males only) are available, ranging from as low as 14 genes (Olds *et al.* 2012) to 552 genes (Previte *et al.* 2014), a remarkably similar estimate to this study. In contrast, the number of differentially expressed genes between BH and HB males was only 74, showing very limited overlap with the differentially expressed gene set in pure F0 males. Combined, prevalent biparental allele-specific expression patterns and higher similarity between F1 males than between pure louse males offer solid evidence of absence of somatic manifestations of PGE in *P. humanus.*

5.4.2. Putative roles for parentally-biased genes under PGE in the human louse

In our ASE gene panel, only 12 (~7%) were found to be expressed in a consistent parent-of-origin manner. To be consistent with terminology commonly employed in ASE literature, I will hereafter refer to these genes as imprinted strictly in the sense of preferential or exclusive expression of one of the parental alleles. However, it is important to note that the term imprinting also refers to the presence of heritable signals differentially marking both alleles, a distinction which is particularly relevant under PGE (Nur 1990; Herrick & Seger 1999). The presence of a fraction of imprinted genes in human lice was anticipated: in the absence of somatic adaptations to prevent or reduce expression of paternal alleles, maternally imprinted genes can be expected to be involved in reproductive functions under PGE, chiefly during spermatogenesis. In

particular, faithful segregation of paternal and maternal chromosomes in meiosis is key to evolution and maintenance of PGE. In louse male meiosis, differential segregation most likely occurs during the last of a series of mitotic divisions preceding the formation of active spermatids (McMeniman & Barker 2005; de la Filia *et al.* 2017). This stage is a clear candidate for PGE-inducing maternal genome adaptations to manifest themselves and control segregation of parental chromosomes. In this context, the finding that a *mcm5* ortholog is among the genes showing expression patterns consistent with maternal imprinting is very suggestive. In *Drosophila* females, *mcm5* mutants have increased rates of chromosomal nondisjunction during meiosis (Lake *et al.* 2007). However, this increase in nondisjunction could be related to failure of *mcm5* mutants to resolve chiasmata formation (Lake *et al.* 2007), so the potential role that *mcm5* and other members of its conserved complex (Forsburg 2004) could play in achiasmatic louse males is unclear.

Another expected category of maternally imprinting genes under PGE are those involved in tagging and expression control of parental alleles. In mealybug and sciarid flies, parental chromosomes display differential patterns of epigenetic signals, such as histone modifications (Goday & Ruiz 2002; Bongiorni & Prantera 2003; Khosla *et al.* 2006; Escribá *et al.* 2011). *Su(z)2*, the *Drosophila* ortholog of the maternally-biased PHUM328760 gene, is an adult-expressed functional homolog of *Psc*, a member of the Polycomb complex PRC1. PRC1 is a multifunctional transcriptional repressor than interacts with H3K27 methylation marks (Lo *et al.* 2009; Schwartz & Pirrotta 2008). H3K27 is a conserved marker of facultative heterochromatin which has been shown to specifically mediate processes of gene repression, such as *hox* silencing in flies or X chromosome inactivation in mammals (Plath *et al.* 2003) and paternal genome silencing in mealybugs (Bain 2018).

Other potential sources of parental imprinting under PGE are not directly linked to exclusion of paternal alleles from male gametes. Haplodiploid systems create scope for evolution for imprinted genes due to the unique relatedness patterns arising from asymmetries in gene inheritance (Haig 1992). Under arrhenotokoky (i.e. haplodiploidy *sensu stricto*), haploid males contribute the same alleles to all their daughters, creating a higher relatedness (3/4) between full sisters (Hamilton 1964). In PGE, all siblings share this higher relatedness since males transmit the same alleles to both sexes

(barring leakage of paternal chromosomes). Therefore, PGE siblings are more related through their paternal genomes that through the maternal, creating differential selective pressures on paternally-inherited alleles (e.g. to be less selfish) and maternally-inherited alleles. On the other hand, polyandry has the opposite effect, since it reduces paternal relatedness between broods (Haig 1992; Queller 2003). Interestingly, in another haplodiploid species, the honeybee, the ortholog of PHUM544500 (a maternally-imprinted cleavage and polyadenylation specificity factor), shows DNA methylation patterns that are consistent with parental imprinting (Remnant *et al.* 2016). However, in human lice, these predictions may not apply, due to a combination of factors: absence of social interactions (although kin live in close proximity), high levels of inbreeding and sib-mating (which increases relatedness between parental alleles) and reduced competition over resources (Perotti *et al.* 2004; Takano-Lee *et al.* 2005; Ascunce *et al.* 2013). Additionally, imprinting has been suggested to arise to maximise expression of alleles inherited from the parent that has experienced higher sex-specific selection pressures on the related trait (Day & Bonduriansky 2004). Resilin, the elastic protein encoded by the only paternally-imprinted gene with a known function found in this study, may facilitate the extreme cuticle enlarging required for blood meals (Benoit *et al.* 2016). Due to their small size compared to females (Busvine 1978), the role of resilin could be particularly important in males.

Finally, an unexpected finding was the high frequency of genes (~20%) exhibiting asymmetric parent-of-origin patterns between reciprocal F1s. Most of these genes showed a maternal bias in BH males but were expressed biparentally in HB. Naively, the opposite pattern could have been expected: leakages of paternal chromosomes during spermatogenesis are more common in body than in head lice (Chapter 4), which would suggest a higher degree of maternal allele expression in head lice. Unequal parent-of-origin expression biases between F1s have also been found in honey bees (Kocher *et al.* 2015), where they were attributed to asymmetric effects of mito-nuclear incompatibilities between parental populations (Turelli & Moyle 2007). In *P. humanus*, the mitochondrial gene is fragmented into 18 minichromosomes that recombine frequently, which may accelerate evolution of mitochondrial genes (Shao *et al.* 2009). Indeed, louse mitochondrial genomes exhibit remarkably rapid evolutionary rates (Johnson *et al.* 2003; Yoshizawa & Johnson 2003). As a result, mito-nuclear mismatches

may be frequent in human lice and account for this subset of asymmetrically-biased genes.

5.4.3. Limitations of this study arise from the paucity of discriminant SNPs

Genome-wide estimates of patterns of parent-of-origin expression rely on sufficient density of fixed SNPs between parental genomes. Although initially a putative set of >5,000 SNPs were found in all F1 transcriptomes, less than 1,000 SNPs were found to be actually fixed between pure head and body louse transcriptomes. Of these, only 450 mapped to known exons and showed homogeneity across reciprocal cross replicates. As a result, only ~1.5% of genes predicted in *P. humanus* could be covered by this limited SNP panel, which falls short of being representative of the entire genome. Typically, ASE studies at intraspecific scale benefit from a higher number of informative SNPs, allowing a broader estimation of parent-of-origin expression across the genome (see for example Babak *et al.* 2008; Wang *et al.* 2011; Kocher *et al.* 2015; Kincaid-Smith *et al.* 2018).

There are several reasons to explain the paucity of fixed SNPs in this study. The aforementioned studies employ distantly related and highly inbred laboratory strains (e.g. Wade & Daly 2005; Clément *et al.* 2013) or different subspecies (Page Jr & Amdam 2007). Head and body lice, however, are very closely related. The specific status of head and body light has been long debated, but mounting evidence supports that their phenotypic differences are due to ecological factors and even their subspecific status has been questioned (Light *et al.* 2008). Currently, they are considered different ecotypes of *P. humanus:* body lice are believed to emerge regularly from head louse populations by colonising new breeding grounds in human clothes (Li *et al.* 2010). Both laboratory colonies used in this study derive from US populations and are not raised under a strict inbreeding regime. A previous study of sequence divergence at the transcriptomic level between the same strains revealed low levels of interecotype nucleotide divergence (0.1-1.3%) that were comparable to intraecotype diversity (0-1.1%) (Olds *et al.* 2012).

Also, the analysis presented here is more conservative in SNP validation within transcriptomes. In particular, I implemented two steps that are absent from the ASE

analysis pipeline this study is based upon (Wang & Clark 2014). First, after DNA-seq based variant calling, I used pure F0 male transcriptomes for additional validation of discriminant sites. Second, I filtered out discordant SNPs between replicates prior to assignment to predicted genes. Since single males were sequenced to obtain F1 transcriptomes, these steps were necessary to avoid obtaining spurious ASE patterns arising from any of the replicates not being heterozygous for the BB and HH bases assigned during variant calling, a likely scenario due to segregating variation within parental ecotypes.

A study purely aimed at identifying imprinted genes could have relaxed these additional SNP validation steps (at the expense of an increase in false negatives), but the main goal of this chapter was to determine the degree of biparental expression in lice, which could be more prone to be confounded by polymorphism. In this regard, the limited number of genes with ASE information is sufficient to confidently determine the mode of gene expression in human lice, yet obtaining a comprehensive catalogue of imprinted genes will require additional sequencing effort. A direct solution to the lack of differentiation between louse ecotypes would be sequencing the parents of all F1 replicates to directly assign parental genotypes at heterozygous sites individually for each male.

5.4.4. Conclusions and future directions

In this chapter, I have shown that adult males of *Pediculus humanus* display biparental gene expression, which constitutes the first known case of a PGE species in which genetic activity of paternal chromosomes in the soma is not affected by embryonic heterochromatinization or (partial or complete) elimination. Additionally, I have identified a preliminary subset of genes exhibiting parent-of-origin specific expression, some of which show a direct involvement in meiosis and gene-specific transcriptional repression in other insect species. The absence of adaptations to silence the paternal genome as a whole in the human louse can facilitate unmasking genes involved in PGE (either enforcing it when maternally imprinted or combating elimination when paternally imprinted), most specifically during spermatogenesis. The analysis presented in this chapter is a necessary first step to confirm biparental expression in human louse from which future studies specifically aimed at identifying such biased genes can stem. As in

Chapter 3, obtaining transcriptomes from germ-line tissues exclusively would be an immediate strategy to screen for imprinted genes directly in the arena of paternal chromosome elimination.

From an evolutionary perspective, the description for the first time of a form of basal PGE in a parasitic louse is an exciting finding. The only other louse species that has been shown to reproduce through PGE is the *Liposcelis* booklouse (Hodson *et al.* 2017), a free living louse. In contrast to *Pediculus*, *Liposcelis* exhibits paternal chromosome heterochromatinization. Free living (order Psocoptera) and parasitic lice (order Phthiraptera) diverged 118-286 Mya (Misof *et al.* 2014), putting a putative timeframe on the emergence of PGE in lice. Although direct evidence in additional species is needed, there are common features between *Pediculus* and species from other suborders of parasitic lice that suggest that PGE is present in the whole group: the same modified spermatogenesis (Tombesi & Papeschi 1993; Tombesi *et al.* 1999) and unusual sperm structure (Dallai & Afzelius 1991; Ross & Normark 2015). Lineages of parasitic lice diverged at least 46-67 Mya (Misof *et al.* 2014). Since no somatic condensation or elimination of chromosomes have been reported in any species, Phthiraptera could represent a long evolutionary period of basal PGE. Although it is premature to draw such conclusions until more data is gathered, parasitic lice could represent a promising group to scrutinise for possible early losses and gains of PGE and expose emerging maternal and paternal adaptations over control of spermatogenesis.

Chapter 6

Concluding remarks

6.1. Thesis overview and summary of main findings

The extraordinary diversity of sexual reproduction is still a fundamental question in evolutionary biology. The interplay between evolutionary forces and processes driving transitions between different modes of reproduction has been, is, and will remain a major field of research. Sexual reproduction has been ripe for an affluence of evolutionary theory, but matching theoretical predictions to empirical evidence is not always straightforward. The main goal of this thesis was to provide empirical support for a range of models on the role of intragenomic conflict in the evolution of a bizarre—and perhaps historically overlooked—genetic system, paternal genome elimination, which dramatically illustrates the complexities of reproductive modes. The scope for conflict between maternal and paternal alleles under PGE is very strong, as exclusion of paternal chromosomes from the germline of males creates a battleground for whole parental genomes over transmission to following generations. Such conflict has been predicted, in formal or verbal models, to be not only a consequence of PGE but the driving force behind the emergence of this genetic system and its different forms across arthropods (Brown 1964; Bull 1979; Haig 1993; Herrick & Seger 1999; Normark 2006; Burt & Trivers 2006; Ross *et al.* 2010a; Normark & Ross 2014; Gardner & Ross 2014). The extent to which these models explain the manifestations and taxonomic distribution of PGE is enormous, but they have suffered from a lack of empirical validation. In this thesis, I have explicitly tested some of the predictions arising from such models directly in the citrus mealybug, *Planococcus citri*, and indirectly in the human louse, *Pediculus humanus.* Also, I have fully characterised the form of PGE that has evolved in the latter, since the apparent polymorphism in transmission of paternal chromosomes between body louse males (McMeniman & Barker 2005) suggested an incipient stage in the evolution of PGE that was hitherto undiscovered in any other species. The goal of the experiments presented here can be reduced to answering two single questions: can paternal genomes escape elimination and transcriptional suppression under PGE?; and, if so, can the existence of an arms race between parental genome be inferred from incomplete PGE?

In **Chapter 2**, this question was explored by tracking down inheritance patterns of microsatellite markers—a traceable genotype—and sex pheromone preferences—a heritable phenotype—through a three-generation mealybug pedigree. Transmission of inheritance markers from fathers to their offspring was studied in both intraspecific and hybrid crosses to determine whether males can transmit paternal alleles when exposed to a divergent maternal background. I reasoned that higher rates of leakages in hybrid crosses would be indicative a historical evolutionary arms race within *P. citri* or the closely related *P. ficus*, while the opposite outcome—more frequent escapes in intraspecific crosses—would reveal ongoing antagonistic coevolution between parental genomes in current populations of *P. citri*. Additionally, the species-specificity of sex pheromone responses would allow to further explore an arms race at the interspecific scale by coupling transmission and expression of a paternal trait. I found occasional instances of transmission of paternal microsatellites alleles through both hybrid and intraspecific males at similar low frequency, but no evidence of transmission of pheromone preferences. These results cannot therefore demonstrate the existence of an arms race. However, the extreme male mortality of males deriving from crosses between *P. ficus* mothers and *P. citri* fathers only allowed to explore paternal escapes through the reciprocal CF hybrids, and the intensity of these asymmetric hybrid incompatibilities—which are not expected under silencing of paternal chromosomes in lecanoid PGE—could constitute indirect evidence for drivers and suppressors of silencing of paternal alleles manifesting in one of the cross directions exclusively.

In **Chapter 3**, I directly examined whether paternal alleles are completely silenced in mealybugs by obtaining genome-wide patterns of gene expression in somatic and reproductive tissues of hybrid mealybug males. I hypothesised that an evolutionary arms race would result in differential patterns of expression of paternal chromosomes between soma and testis. I used a combination of genomic and transcriptomic data to identify informative and discriminant SNPs between the genomes of *P. citri* and *P. ficus* and obtained allele counts at these sites from the hybrid transcriptomes. I found that only less than a third of genes exhibit complete silencing in the soma, in contrast to 80% in testis. Many of the genes that exhibited biallelic expression were assigned to fundamental cell processes, such as mitochondrial functions and lipid metabolism, suggesting that paternal chromosomes are reactivated to contribute to these processes

in a tissue-specific manner. The stronger suppression of paternal allele expression in germline tissues could indicate maternal adaptations to prevent the expression of putative anti-PGE responses, yet I did not identify any candidate genes which could be inferred to be involved in gametogenesis.

In **Chapter 4**, I determined patterns of microsatellite allele inheritance through males from both ecotypes of *P. humanus*. I found that all males of head and body lice have transmission patterns consistent with PGE, which is therefore not polymorphic in this species. However, I also found relatively high rates of paternal chromosome leakages, particularly in body lice, which reveals very strong scope for intragenomic conflict in *P. humanus*. From a more applied perspective, in this chapter I also argued that the genetic system of *P. humanus* should be considered when designing treatment strategies to counteract the current crisis in human louse control derived from increased resistance to available pediculicides.

Finally, having confirmed the occurrence of PGE in *P. humanus*, in **Chapter 5** I determined whether paternal chromosomes are expressed in the human louse soma, as was suggested by the lack of cytogenetic observations. I employed the same approach developed in Chapter 3 to identify discriminant SNPs between head and body louse genomes and estimate proportions of maternal and paternal alleles in transcriptomes of individual males deriving from interecotype crosses. Although hampered by a limited number of informative SNPs, I found that expression is predominantly biparental across the genome. However, I identified a fraction of genes exhibiting allelic expression patterns consistent with *cis*-regulatory divergence and, most importantly in the context of PGE, parent-of-origin expression. Functional interrogation of such genes with preferential maternal expression revealed a putative involvement in meiotic segregation and gene-specific epigenetic silencing, thus constituting promising candidates to identify maternally-expressed genes responsible for PGE in males.

6.2. Final thoughts and future directions

6.2.1. Can paternal genomes challenge PGE?

In Chapters 2 and 4, I have explored whether paternal chromosomes can escape elimination in germline PGE. The ability of paternal chromosomes to become incorporated into active sperm is a major assumption of the arms race hypothesis of transitions between forms of PGE and reversions from PGE to diplodiploidy (Herrick & Seger 1999; Burt & Trivers 2006; Ross *et al.* 2010a). Paternal escapes are difficult to detect and study due to their infrequent occurrence and the low sensitivity and high error rates of methodologies used to uncover them. For example, possible events of paternal transmission in a germline PGE species, the coffee borer beetle *Hypothenemus hampei* could not be distinguished from misclassification of individuals exhibiting genotypes incompatible with PGE or were dismissed (Borsa & Kjellberg 1996; Borsa & Coustau 1996). Other findings of paternal chromosome incorporation into sperm were dependent on damaging experimental alterations (Nur 1970), and therefore more interesting from a mechanistic point of view than representative of natural populations. These two chapters build upon previous reports of paternal chromosome transmission in human lice and mealybugs (McMeniman & Barker 2005; Kol-Maimon *et al.* 2014a), which were hampered by limitations of their experimental toolkit. I detected paternal escapes in both intraspecific and hybrid *Planococcus* males and through head and body louse males, although an evolutionary arms race cannot be inferred from these results, as the data does not support this hypothesis (in mealybugs) or the experimental design did not intend to explicitly demonstrate such hypothesis (in lice). A comparison between paternal leakages in mealybugs and lice brings out some interesting, if possibly premature, considerations. First, the mode of paternal escape appears to be the same in both species: replacement of maternal homologs, rather than complete substitution of the whole maternal complement or incorporation of paternal chromosomes to sperm in addition to the maternal set. Second, human lice and mealybugs both share basal forms of PGE, where paternal chromosomes are not destroyed prior to spermatogenesis and undergo meiosis with their maternal counterparts, unlike in more derived forms of PGE (Ross *et al.*, 2010a), thus potentially creating more scope for paternal resistance adaptations to evolve. It is tempting to associate the higher rates of leakage in *P. humanus*, particularly in body lice, to the lack of global silencing of paternal chromosomes, which allows for anti-PGE paternal adaptations to manifest themselves

without depending on reversal of heterochromatization. However, these observations remain anecdotal until more cases of incomplete PGE are reported and evidence can be drawn from their phylogenetic distribution and comparisons between the different manifestations of this genetic system.

Additional experimental work will be needed to fully understand how paternal chromosomes escape, what is the distribution of paternal escapes through the entire paternal genome and whether there is a heritable component to paternal chromosome leakages. Particular effort should be devoted to reassessing escapes within *P. citri*, due to the limited number of polymorphic markers for intraspecific crosses (only three loci, often homozygous in F1 males) and the lack of unique alleles to the parental lines. As discussed in section 2.5.1, a genotype-by-sequencing approach to detecting escapes and inferring their genomic distribution would be a superior approach to the one implemented here. The availability of a reference genome for *P. citri*—to which the data generated for Chapter 3 was instrumental—and *P. humanus* will facilitate this effort. Beyond these two species, other germline PGE species with or without paternal chromosome silencing that can be easily bred in the laboratory, such as the coffee borer beetle, booklice or sciarid flies, are promising candidates for future inheritance studies aimed at determining whether paternal chromosome escapes that could challenge maternal control under PGE are the norm rather than the exception.

A major contribution of this thesis is the direct exploration of patterns of paternal genome expression at a genome-wide scale in mealybugs and human lice. As I have discussed along this thesis, residual transcriptional activity of paternal chromosomes under lecanoid PGE had been long suspected (Nelson-Rees 1962; Nur & Chandra 1963; Brown & Nur 1964; Nur 1966; Nur 1967; Nur 1990; Ross *et al.* 2010a), but Chapter 3 is the first direct empirical demonstration of paternal chromosome expression. The implications of paternal chromosome reactivation in mealybugs are important, not only directly in the context of an arms race between parental genomes—which the data in Chapter 3 supports—but also because it offers a plausible explanation for the seemingly paradoxical hybrid incompatibilities in ♀*P. ficus* x ♂*P. citri* hybrid males and other mealybug hybrids (Nur & Chandra 1963). If the activity of paternal chromosomes is required to complement the maternal genome, as the functional profiling of non-silenced genes seems to indicate, the lack of viability of these males is most likely related to failure

in recruiting paternal alleles, as hybrid females escape such incompatibilities. Of course, experimentally confirming this hypothesis would require examining allele-specific expression patterns in hybrid males that do not survive—which presents an obvious difficulty. The magnitude of hybrid incompatibilities in FC seems to be dependent on experimental populations, as viability of adult hybrids vary among this thesis and other studies (Rotundo & Tremblay 1982; Kol-Maimon *et al.* 2014a), so it should be possible to conduct these crosses with other genetic backgrounds and examine differences between reciprocal hybrids. Another immediate follow-up to Chapter 3 would be replicating this experiment at the intraspecific scale, in order to disentangle the functional significance of paternal chromosome reactivation from possible hybrid effects in CF males. Other future directions, as discussed in section 3.5, would be to sequence germline tissues of earlier larval stages and expose patterns of parental allele expression during the peak of spermatogenesis, as well as dissecting tissue-specific differences in the soma to gain an understanding of the temporal and spatial regulation of PGE silencing.

The main goal of Chapter 5, on the other hand, was more descriptive. The finding that both parental genomes are equally expressed, barring a subset of imprinted genes, positions *P. humanus* as the PGE species showing the most basal form of this genetic system. The other two known germline PGE groups which seem to lack adaptations to silence paternal chromosomes in the soma, Sciaridae and Cecidomyiidae, still eliminate paternally-inherited X-chromosomes early in development (Goday & Esteban 2001; Benatti *et al.* 2010), so that human lice are to current knowledge the sole PGE species with complete diploid expression. One need not emphasise the practical difficulties of working with human lice, which require a constant supply of human blood and a laborious *in vitro* rearing system (Takano-Lee *et al.* 2003), but their exceptionality as a study system for intragenomic conflict between parental alleles in the absence of heterochromatization is unquestionable. Hopefully, PGE will be confirmed in more louse species in the near future. It is particularly exciting that only two years have separated the full description of PGE in the booklouse *Liposcelis* sp. (Hodson *et al.* 2017) and *P. humanus*, highlighting the potential of Psocodea to understand the evolution of this genetic system.

6.2.2. Wider implications: insect genomic imprinting and applied perspectives

Of perhaps broader interest to evolutionary biologists who are unfamiliar with PGE, a chief result of Chapters 3 and 5 is the demonstration of parent-of-origin-dependent gene expression in mealybugs, where it affects the majority of the genome, and human lice, where it appears to occur in a gene-by-gene basis. Recent years have seen an increased interest in genomic imprinting beyond mammals and flowering plants (de la Casa-Esperón 2012; MacDonald 2012). Although our understanding of genomic imprinting in insects is still in its infancy, this thesis can warn against two obstacles that could hinder our progress in this field. First, the premature conclusion that gene-by-gene imprinting should be extremely rare in insects, a view perhaps excessively driven by the apparent lack of natively imprinted genes in adult *Drosophila* (Menon & Meller 2010; Coolon *et al.* 2012). Second, an almost exclusive focus on eusocial Hymenoptera to search for parent-of-origin gene expression, due to a too narrow reading of the kinship theory of genomic imprinting. Although this approach is justified, given the high scope for interactions between relatives with asymmetric genetic relationships under eusociality (Queller 2003; Patten *et al.* 2014), it can lead to extreme conclusions (Wang *et al.* 2016). It has long been an unfortunate tendency to fully conflate haplodiploidy and eusociality as two sides of the same coin, which has led to overlooking the impact of the same genetic relatedness asymmetries in non-social species whose life histories promote strong associations between kin—such as most PGE taxa. Both obstacles can be overcome with a more diverse sampling of insect species, which will no doubt be available before long.

Hopefully, another useful contribution of this thesis, even though in discussion only, will be stressing the potential significance of PGE for the evolution of insecticide resistance. It is surprising that the initial report of such a dramatic transmission ratio distortion in body lice (McMeniman & Barker 2005) did not merit consideration in the context of the rapid evolution of resistance to pediculicides in current human louse populations, to which extensive research has been devoted since (Clark 2018). According to the expression patterns found in Chapter 5, the existing models for faster evolution of resistance in haplodiploids (Crozier 1985; Havron *et al.* 1987; Caprio & Hoy 1995; Brun *et al.* 1995; Denholm *et al.* 1998) will need to be revised for *P. humanus* to determine whether the asymmetric transmission patterns through males and females

can influence *per se* the spread of resistance alleles. At this stage, without formal models or empirical evidence in this direction, this possibility should be considered speculative only. But the potential consequences of PGE and other asymmetric systems for inheritance and expression of traits—see Appendix 1 (de la Filia *et al.* 2015) as an example—demand taking into account the diversity of reproductive modes for these and other practical challenges. Positively, in future years, a better understanding of the evolutionary causes and consequences of diversity in sexual reproduction will go in hand with an increased recognition of its importance in applied biology.

6.3.3. Coda

To understand a huge range of biological phenomena, it is often useful to study dramatic deviations from the norm to learn how widely conserved rules can be twisted to produce such anomalies. A more complete understanding of both the molecular and ecological features and the evolutionary history of PGE and other bizarre genetic systems can shed light on the intricate balance between the different parties that compose a eukaryotic genome and their power games. Theory usually leads the way, but comparative and experimental validation must follow suite. In this thesis, I have attempted to broaden our knowledge of PGE and its evolutionary foundations by shortening the distance between predicted and demonstrable. The work presented here was built upon the momentous contribution of many brilliant and creative empiricists that managed to make great progress in understanding such a strange mode of reproduction with a much more limited set of tools than the ones at our current disposal. Having been able to explore some of their hypotheses and prove them right gives me an enormous satisfaction. Empirical studies on PGE lived a golden age in the mid-20th century, yet somehow this fascinating system fell into relative obscurity while waiting for new generations of scientists to pick up the torch in the genomics era. It is my hope that this work will contribute to this effort.

References

Adler, P.N., Charlton, J. & Brunk, B., 1989. Genetic interactions of the suppressor 2 of zeste region genes. *Developmental Genetics*, 10(3), pp. 249–260.

Andersen, S.O., 2010. Studies on resilin-like gene products in insects. *Insect Biochemistry and Molecular Biology*, 40(7), pp. 541–551.

Ascunce, M.S. *et al.*, 2013. Nuclear genetic diversity in human lice (*Pediculus humanus*) reveals continental differences and high inbreeding among worldwide populations. *PLoS One*, 8(2), p. e57619.

Ashman, T. L. *et al.*, 2014. Tree of Sex: A database of sexual systems. *Scientific Data*, 1, p. 140015.

Avise, J.C. and Mank, J.E., 2009. Evolutionary perspectives on hermaphroditism in fishes. *Sexual Development*, 3(2-3), pp.152-163.

Babak, T. *et al.*, 2008. Global survey of genomic imprinting by transcriptome sequencing. *Current Biology*, 18(22), pp. 1735–1741.

Bachtrog, D. *et al.*, 2014. Sex determination: why so many ways of doing it? *PLoS Biology*, 12(7), p. e1001899.

Bacot, A., 1917. A contribution to the bionomics of Pediculus humanus (vestimenti) and Pediculus capitis. *Parasitology*, 9(2), pp. 228–258.

Barrett, S.C.H., 2010. Understanding plant reproductive diversity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1537), pp. 99–109.

Bates, D. *et al.*, 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software; Vol 1, Issue 1 (2015)*, 67(1), pp. 1–48.

Bayreuther, K., 1955. Holokinetische Chromosomen bei *Haematopinus suis* (Anoplura, Haematopinidae). *Chromosoma*, 7(1), pp. 260–270.

Beckmann, R. & Herrmann, J.M., 2015. Mitoribosome oddities. *Science*, 348(6232), pp. 288–289.

Bell, G.D.M. *et al.*, 2013. RNA-seq analysis of allele-specific expression, hybrid effects, and regulatory divergence in hybrids compared with their parents from natural populations. *Genome Biology and Evolution*, 5(7), pp. 1309–1323.

Benatti, T.R. *et al.*, 2010. A neo-sex chromosome that drives postzygotic sex determination in the Hessian fly (*Mayetiola destructor*). *Genetics*, 184(3), pp. 769–777.

Benoit, J.B. *et al.*, 2016. Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nature Communications*, 7, ncomms10165.

Berlowitz, L., 1965. Correlation of genetic activity, heterochromatization, and RNA metabolism. *Proceedings of the National Academy of Sciences*, 53(1), pp. 68–73.

Berlowitz, L., Loewus, M.W. & Pallotta, D., 1968. Heterochromatin and genetic activity in mealy

bugs I. Compensation for inactive chromatin by increase in cell number. *Genetics*, 60(1), pp. 93–99.

Beukeboom, L. & Perrin, N., 2014. The evolution of sex determination. Oxford University Press, USA.

Bierl-Leonhardt, B.A. *et al.*, 1981. Isolation, identification and synthesis of the sex pheromone of the citrus mealybug, *Planococcus citri* (Risso). *Tetrahedron Letters*, 22(5), pp. 389–392.

Billiard, S. *et al.*, 2012. Sex, outcrossing and mating types: unsolved questions in fungi and beyond. *Journal of Evolutionary Biology*, 25(6), pp. 1020–1038.

Birky, C.W., 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences*, 92(25), pp. 11331–11338.

Blackmon, H., Hardy, N.B. & Ross, L., 2015. The evolutionary dynamics of haplodiploidy: genome architecture and haploid viability. *Evolution*, 69(11), pp. 2971–2978.

Blackmon, H., Ross, L. & Bachtrog, D., 2016. Sex determination, sex chromosomes, and karyotype evolution in insects. *Journal of Heredity*, 108(1), pp. 78–93.

Bolger, A.M., Lohse, M. & Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), pp. 2114-2120.

Bongiorni, S., Cintio, O. & Prantera, G., 1999. The relationship between DNA methylation and chromosome imprinting in the coccid *Planococcus citri*. *Genetics*, 151(4), pp. 1471–1478.

Bongiorni, S. *et al.*, 2001. Facultative heterochromatization in parahaploid male mealybugs: involvement of a heterochromatin-associated protein. *Development*, 128(19), pp. 3809–3817.

Bongiorni, S. & Prantera, G., 2003. Imprinted facultative heterochromatization in mealybugs. *Genetica*, 117(2–3), pp. 271–279.

Bongiorni, S. *et al.*, 2004. Inverted meiosis and meiotic drive in mealybugs. *Chromosoma*, 112(7), pp. 331–341.

Bongiorni, S. *et al.*, 2007. Epigenetic regulation of facultative heterochromatinisation in *Planococcus citri* via the Me (3) K9H3-HP1-Me (3) K20H4 pathway. *Journal of Cell Science*, 120(6), pp. 1072–1080.

Bongiorni, S. *et al.*, 2009. Epigenetic marks for chromosome imprinting during spermatogenesis in coccids. *Chromosoma*, 118(4), pp. 501–512.

Borsa, P. & Coustau, C., 1996. Single-stranded DNA conformation polymorphism at the Rdl locus in *Hypothenemus hampei* (Coleoptera: Scolytidae). *Heredity*, 76(2), pp. 124–129.

Borsa, P. & Kjellberg, F., 1996. Experimental evidence for pseudo-arrhenotoky in *Hypothenemus hampei* (Coleoptera: Scolytidae). *Heredity*, 76, pp. 130–135.

Boulton, R.A., Collins, L.A. & Shuker, D.M., 2015. Beyond sex allocation: the role of mating systems in sexual selection in parasitoid wasps. *Biological Reviews*, 90(2), pp. 599–627.

Bressa, M.J., Papeschi, A.G. & Toloza, A.C., 2015. Cytogenetic features of human head and

body lice (Phthiraptera: Pediculidae). *Journal of Medical Entomology*, 52(5), pp. 918–924.

Brown, S.W., 1963. The Comstockiella system of chromosome behavior in the armored scale insects (Coccoidea: Diaspididae). *Chromosoma*, 14(4), pp. 360–406.

Brown, S.W., 1964. Automatic frequency response in the evolution of male haploidy and other coccid chromosome systems. *Genetics*, 49(5), pp. 797–817.

Brown, S.W., 1967. Chromosome systems of the Eriococcidae (Coccoidea-Homoptera). *Chromosoma*, 22(2), pp. 126–150.

Brown, S.W., 1972. Developmental control of heterochromatinization in coccids. *The meiotic process: pairing, recombination, and chromosome movements*, 4, p. 130.

Brown, S.W. & Nelson-Rees, W.A., 1961. Radiation analysis of a lecanoid genetic system. *Genetics*, 46(8), p.983.

Brown, S.W. & Nur, U., 1964. Heterochromatic chromosomes in the coccids. *Science*, 145(3628), pp. 130–136.

Brown, S.W. & Wiegmann, L.I., 1969. Cytogenetics of the mealybug *Planococcus citri* (Risso) (Homoptera: Coccoidea): genetic markers, lethals, and chromosome rearrangements. *Chromosoma*, 28(3), pp. 255–279.

Brun, L.O. *et al.*, 1995. Functional haplodiploidy: a mechanism for the spread of insecticide resistance in an important international insect pest. *Proceedings of the National Academy of Sciences*, 92(21), pp. 9861–9865.

Buglia, G.L. & Ferraro, M., 2004. Germline cyst development and imprinting in male mealybug *Planococcus citri*. *Chromosoma*, 113(6), pp. 284–294.

Bull, J.J., 1979. An advantage for the evolution of male haploidy and systems with similar genetic transmission. *Heredity*, 43(3), pp. 361–381.

Burgess, I.F., 2004. Human lice and their control. *Annual Review of Entomology*, 49(1), pp. 457–481.

Burt, A. & Trivers, R., 1998. Genetic conflicts in genomic imprinting. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1413), pp. 2393–2397.

Burt, A. & Trivers, R., 2006. Genes in conflict: the biology of selfish genetic elements, Harvard University Press, USA.

Busvine, J.R., 1978. Evidence from double infestations for the specific status of human head lice and body lice (Anoplura). *Systematic Entomology*, 3(1), pp. 1–8.

Camacho, J.P.M., Sharbel, T.F. & Beukeboom, L.W., 2000. B-chromosome evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 355(1394), pp. 163–178.

Caprio, M.A. & Hoy, M.A., 1995. Premating isolation in a simulation model generates frequency-dependent selection and alters establishment rates of resistant natural enemies. *Journal of Economic Entomology*, 88(2), pp. 205–212.

Carrel, L. & Willard, H.F., 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*, 434(7031), p. 400.

Carrière, Y., 2003. Haplodiploidy, sex, and the evolution of pesticide resistance. *Journal of Economic Entomology*, 96(6), pp. 1626–1640.

Chamberlain, A.J. *et al.*, 2015. Extensive variation between tissues in allele specific expression in an outbred mammal. *BMC Genomics*, 16(1), p. 993.

Chandra, H.S., 1962. Inverse meiosis in triploid females of the mealy bug, *Planococcus citri*. *Genetics*, 47(10), pp. 1441–1454.

Chandra, H.S., 1963. Cytogenetic studies following high dosage paternal irradiation in the mealy bug, *Planococcus citri*. *Chromosoma*, 14(3), pp. 310–329.

Charlesworth, D. & Charlesworth, B., 1995. Quantitative genetics in plants: the effect of the breeding system on genetic variability. *Evolution*, 49(5), pp. 911–920.

Charnov, E.L., 1979. Simultaneous hermaphroditism and sexual selection. *Proceedings of the National Academy of Sciences*, 76(5), pp. 2480–2484.

Chen, S. *et al.*, 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *bioRxiv*, p. 274100.

Chen, Z. *et al.*, 2016. Global assessment of imprinted gene expression in the bovine conceptus by next generation sequencing. *Epigenetics*, 11(7), pp. 501–516.

Chmátal, L. *et al.*, 2014. Centromere strength provides the cell biological basis for meiotic drive and karyotype evolution in mice. *Current Biology*, 24(19), pp. 2295–2300.

Clark, J.M., 2018. New developments in the control of human lice. *Ectoparasites: Drug Discovery Against Moving Targets*, pp. 119–137.

Clark, J.M. *et al.*, 2013. Human lice: Past, present and future control. *Pesticide Biochemistry and Physiology*, 106(3), pp. 162–171.

Clark, J.M. *et al.*, 2015. Utilization of the human louse genome to study insecticide resistance and innate immune response. *Pesticide Biochemistry and Physiology*, 120, pp. 125–132.

Clément, J.A.J. *et al.*, 2013. Private selective sweeps identified from next-generation pool-sequencing reveal convergent pathways under selection in two inbred *Schistosoma mansoni* strains. *PLoS Neglected Tropical Diseases*, 7(12), p. e2591.

Connelly, J.C. & Leach, D.R.F., 2002. Tethering on the brink: the evolutionarily conserved Mre11–Rad50 complex. *Trends in Biochemical Sciences*, 27(8), pp. 410–418.

Coolon, J.D. *et al.*, 2012. Genomic imprinting absent in *Drosophila melanogaster* adult females. *Cell Reports*, 2(1), pp. 69–75.

Cosmides, L.M. & Tooby, J., 1981. Cytoplasmic inheritance and intragenomic conflict. *Journal of Theoretical Biology*, 89(83), p. 129.

Crespi, B. & Summers, K., 2005. Evolutionary biology of cancer. *Trends in Ecology & Evolution*, 20(10), pp. 545–552.

Crouse, H. V, 1960. The controlling element in sex chromosome behavior in *Sciara*. *Genetics*, 45(10), pp. 1429–1443.

Crowley, J.J. *et al*., 2015. Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. *Nature Genetics*, 47(4), p.353.

Crozier, R.H., 1985. Adaptive consequences of male-haploidy. In Spider mites: Their biology, natural enemies and control. Elsevier, The Netherlands, pp. 201–222.

Culpepper, G.H., 1944. The rearing and maintenance of a laboratory colony of the body louse. *American Journal of Tropical Medicine*, 24(5), pp. 327–329.

Dakin, E.E. & Avise, J.C., 2004. Microsatellite null alleles in parentage analysis. *Heredity*, 93(5), p. 504.

Dallai, R. & Afzelius, B.A., 1991. Sperm flagellum of insects belonging to orders Psocoptera, Mallophaga and Anoplura. Ultrastructural and phylogenetic aspects. *Italian Journal of Zoology*, 58(3), pp. 211–216.

Damon, A., 2000. A review of the biology and control of the coffee berry borer, *Hypothenemus hampei* (Coleoptera: Scolytidae). *Bulletin of Entomological Research*, 90(06), pp. 453–465.

Dasari, V. *et al*., 2018. Epigenetic factors Polycomb (Pc) and Suppressor of zeste (Su(z)2) negatively regulate longevity in *Drosophila melanogaster*. *Biogerontology*, 19(1), pp. 33–45.

Da Silva, E.B., Mendel, Z. & Franco, J.C., 2010. Can facultative parthenogenesis occur in biparental mealybug species? *Phytoparasitica*, 38(1), pp. 19–21.

Dawkins, R., 1976. The selfish gene. Oxford University Press, USA

Day, T. & Bonduriansky, R., 2004. Intralocus sexual conflict can drive the evolution of genomic imprinting. *Genetics*, 167(4), pp. 1537–1546.

Deakin, J.E. *et al*., 2009. Unravelling the evolutionary origins of X chromosome inactivation in mammals: insights from marsupials and monotremes. *Chromosome Research*, 17(5), pp. 671–685.

De la Casa-Esperón, E., 2012. Nonmammalian parent-of-origin effects. In *Genomic Imprinting*, pp. 277–294. Springer, USA.

De la Filia, A.G., Bain, S.A. & Ross, L., 2015. Haplodiploidy and the reproductive ecology of Arthropods. *Current Opinion in Insect Science*, 9, pp. 36–43.

De la Filia, A.G. *et al*., 2017. The unusual reproductive system of head and body lice (*Pediculus humanus*). *Medical and Veterinary Entomology*, 32, pp. 226–234.

Denholm, I. *et al*., 1998. Challenges with managing insecticide resistance in agricultural pests, exemplisfied by the whitefly *Bemisia tabaci. Philosophical Transactions of the Royal Society B: Biological Sciences*, 353(1376), pp. 1757–1767.

Desmond, E. *et al*., 2011. On the last common ancestor and early evolution of eukaryotes: reconstructing the history of mitochondrial ribosomes. *Research in Microbiology*, 162(1), pp. 53–70.

DeVeale, B., Van der Kooy, D. & Babak, T., 2012. Critical evaluation of imprinted gene expression by RNA–Seq: a new perspective. *PLoS Genetics*, 8(3), p. e1002600.

Dobin, A. *et al.*, 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), pp. 15–21.

Durand, R. *et al.*, 2012. Insecticide resistance in head lice: clinical, parasitological and genetic aspects. *Clinical Microbiology and Infection*, 18(4), pp. 338–344.

Durden, L.A. & Musser, G.G., 1994. The sucking lice (Insecta, Anoplura) of the world: a taxonomic checklist with records of mammalian hosts and geographical distributions. *Bulletin of the American Museum of Natural History*, 218.

Ellegren, H., 2000. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in genetics*, 16(12), pp. 551-558.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., *et al.,* 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6: e19379.

Escribá, M.C., Giardini, M.C. & Goday, C., 2011. Histone H3 phosphorylation and non-disjunction of the maternal X chromosome during male meiosis in sciarid flies. *Journal of Cell Science*, 124(10), pp. 1715–1725.

Espírito-Santo, M.M. & Fernandes, G.W., 2007. How many species of gall-inducing insects are there on earth, and where are they? *Annals of the Entomological Society of America*, 100(2), pp. 95–99.

Esteban, M.R. *et al.*, 1997. Role of microtubules and microtubule organizing centers on meiotic chromosome elimination in Sciara ocellaris. *Journal of Cell Science*, 110(6), pp. 721–730.

Ferguson-Smith, A.C., 2011. Genomic imprinting: the emergence of an epigenetic paradigm. *Nature Reviews Genetics*, 12(8), p.565.

Filippova, G.N. *et al.*, 2005. Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Developmental Cell*, 8(1), pp. 31–42.

Fishman, L. & Willis, J.H., 2005. A novel meiotic drive locus almost completely distorts segregation in *Mimulus* (monkeyflower) hybrids. *Genetics*, 169(1), pp. 347–353.

Forsburg, S.L., 2004. Eukaryotic MCM proteins: beyond replication initiation. *Microbiology and Molecular Biology Reviews*, 68(1), pp. 109–131.

Fowden, A.L. *et al.*, 2006. Imprinted genes, placental development and fetal growth. *Hormone Research in Paediatrics*, 65(Suppl. 3), pp. 50–58.

Franco, J.C., Zada, A. & Mendel, Z., 2009. Novel approaches for the management of mealybug pests. In Biorational Control of Arthropod Pests, pp. 233-278. Springer, USA.

Frank, S.A., 1991. Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. *Evolution*, pp. 262–267.

Gardner, A. & Ross, L., 2014. Mating ecology explains patterns of genome elimination. *Ecology*

Letters, 17(12), pp. 1602–1612.

Gardner, A. & Úbeda, F., 2017. The meaning of intragenomic conflict. *Nature Ecology & Evolution*, 1(12), p. 1807.

Garrison, E. & Marth, G., 2012. Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907*.

Gavrilov, I.A., 2007. A catalog of chromosome numbers and genetic systems of scale insects (Homoptera: Coccinea) of the world. *Israel Journal of Entomology*, 37, pp. 1–45.

Gehring, M., Missirian, V. & Henikoff, S., 2011. Genomic analysis of parent-of-origin allelic expression in Arabidopsis thaliana seeds. *PLoS One*, 6(8), p.e23687.

Gill, R.J. & Kosztarab, M., 1997. Pest status of soft scale insects: 3.1. 1. Economic importance. In Soft scale insects–their biology, natural enemies and control, 7, pp. 161-163. Elsevier, The Netherlands.

Goday, C. & Esteban, M.R., 2001. Chromosome elimination in sciarid flies. *Bioessays*, 23(3), pp. 242–250.

Goday, C. & Ruiz, M.F., 2002. Differential acetylation of histones H3 and H4 in paternal and maternal germline chromosomes during development of sciarid flies. *Journal of Cell Science*, 115(24), pp. 4765–4775.

Gould, F. *et al.*, 2010. Sexual isolation of male moths explained by a single pheromone response QTL containing four receptor genes. *Proceedings of the National Academy of Sciences*, 107(19), pp. 8660–8665.

Greciano, P.G. & Goday, C., 2006. Methylation of histone H3 at Lys4 differs between paternal and maternal chromosomes in *Sciara ocellaris* germline development. *Journal of Cell Science*, 119(22), pp. 4667–4677.

Gregg, C. *et al.*, 2010. Sex-specific parent-of-origin allelic expression in the mouse brain. *Science*, 329(5992), pp. 682–685.

GTEx Consortium, 2015. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, 348(6235), pp. 648–660.

Guichoux, E. *et al.*, 2011. Current trends in microsatellite genotyping. *Molecular Ecology Resources*, 11(4), pp. 591–611.

Ghiselin, M.T., 1969. The evolution of hermaphroditism among animals. *The Quarterly Review of Biology*, 44(2), pp.189-208.

Gullan, P.J. & Cook, L.G., 2007. Phylogeny and higher classification of the scale insects (Hemiptera: Sternorrhyncha: Coccoidea). *Zootaxa*, 1668(1), pp. 413–425.

Gullan, P.J. & Kosztarab, M., 1997. Adaptations in scale insects. *Annual Review of Entomology*, 42(1), pp. 23–50.

Haig, D., 1992. Intragenomic conflict and the evolution of eusociality. *Journal of Theoretical Biology*, 156(3), pp. 401–403.

Haig, D., 1993. The evolution of unusual chromosomal systems in coccoids: extraordinary sex ratios revisited. *Journal of Evolutionary Biology*, 6(1), pp. 69–77.

Haig, D., 2000. The kinship theory of genomic imprinting. *Annual Review of Ecology and Systematics*, 31, pp. 9–32.

Hamada, H. *et al.*, 2016. Allele-specific methylome and transcriptome analysis reveals widespread imprinting in the human placenta. *The American Journal of Human Genetics*, 99(5), pp. 1045–1058.

Hamilton, W.D., 1964. The genetical evolution of social behaviour. II. *Journal of Theoretical Biology*, 7(1), pp. 17–52.

Hamilton, W.D., 1967. Extraordinary sex ratios. *Science*, 156(3774), pp. 477–488.

Hansen, R.C. & O'Haver, J., 2004. Economic considerations associated with *Pediculus humanus capitis* infestation. *Clinical Pediatrics*, 43(6), pp. 523–527.

Hartl, D.L. & Brown, S.W., 1970. The origin of male haploid genetic systems and their expected sex ratio. *Theoretical Population Biology*, 1(2), pp. 165–190.

Havron, A. *et al.*, 1987. Selection on the male hemizygous genotype in arrhenotokous insects and mites. *Entomophaga*, 32(3), pp. 261–268.

Heap, G.A. *et al.*, 2009. Genome-wide analysis of allelic expression imbalance in human primary cells by high-throughput transcriptome resequencing. *Human Molecular Genetics*, 19(1), pp. 122–134.

Heard, E. & Disteche, C.M., 2006. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes & Development*, 20(14), pp. 1848–1867.

Hedrick, P.W. & Parker, J.D., 1997. Evolutionary genetics and genetic variation of haplodiploids and X-linked genes. *Annual Review of Ecology and Systematics*, pp. 55–83.

Henter, H.J., 2003. Inbreeding depression and haplodiploidy: experimental measures in a parasitoid and comparisons across diploid and haplodiploid insect taxa. *Evolution*, 57(8), pp. 1712–1793.

Herrick, G. & Seger, J., 1999. Imprinting and paternal genome elimination in insects. In Genomic Imprinting, pp. 41–71. Springer, USA.

Hindle, E. & Pontecorvo, G., 1942. Mitotic divisions following meiosis in *Pediculus corporis* males. *Nature*, 149, p.668.

Hinkens, D.M., McElfresh, J.S. & Millar, J.G., 2001. Identification and synthesis of the sex pheromone of the vine mealybug, *Planococcus ficus*. *Tetrahedron Letters*, 42(9), pp. 1619–1621.

Hodson, C.N. *et al.*, 2017. Paternal genome elimination in *Liposcelis* booklice (Insecta: Psocodea). *Genetics*, 206(2), pp. 1091–1100.

Hoff, K.J. *et al.*, 2015. BRAKER1: unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*, 32(5), pp. 767–769.

Hoffman, J.I. & Amos, W., 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology*, 14(2), pp. 599–612.

Hoffmann, A.A. *et al.*, 2008. A high incidence of parthenogenesis in agricultural pests. *Proceedings of the Royal Society of London B: Biological Sciences*, 275(1650), pp. 2473–2481.

Hothorn, T., Bretz, F. & Westfall, P., 2008. Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3), pp. 346–363.

Hughes-Schrader, S., 1948. Cytology of coccids (Coccoidea-Homoptera). In Advances in Genetics, 2, pp. 127–203. Academic Press.

Hughes-Schrader, S. & Ris, H., 1941. The diffuse spindle attachment of coccids, verified by the mitotic behavior of induced chromosome fragments. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 87(3), pp. 429–456.

Hurst, L.D., 1992. Intragenomic conflict as an evolutionary force. *Proceedings of the Royal Society of London B: Biological Sciences*, 248(1322), pp. 135–140.

Hurst, L., 1995. Selfish genetic elements and their role in evolution: the evolution of sex and some of what that entails. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 349(1329), pp. 321–332.

Hurst, L.D. & Hamilton, W.D., 1992. Cytoplasmic fusion and the nature of sexes. *Proceedings of the Royal Society of London B: Biological Sciences*, 247(1320), pp. 189–194.

Hurst, L.D. & Pomiankowski, A., 1991. Causes of sex ratio bias may account for unisexual sterility in hybrids: a new explanation of Haldane's rule and related phenomena. *Genetics*, 128(4), pp. 841–858.

Hurst, G.D.D. & Werren, J.H., 2001. The role of selfish genetic elements in eukaryotic evolution. *Nature Reviews Genetics*, 2(8), pp. 597–606.

Husnik, F. *et al.*, 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell*, 153(7), pp. 1567–1578.

Jaenike, J., 2001. Sex chromosome meiotic drive. *Annual Review of Ecology and Systematics*, 32(1), pp. 25–49.

Johnson, K.P. *et al.*, 2003. Dramatically elevated rate of mitochondrial substitution in lice (Insecta: Phthiraptera). *Molecular Phylogenetics and Evolution*, 26(2), pp. 231–242.

Johnson, N.A., 2010. Hybrid incompatibility genes: remnants of a genomic battlefield? *Trends in Genetics : TIG*, 26(7), pp. 317–325.

Jones, S.F. & Infante, J.R., 2015. Molecular pathways: fatty acid synthase. *Clinical Cancer Research*, 21(24), pp. 5434–5438.

Kaiser, S.A., Taylor, S.A., Chen, N., Sillett, T.S., Bondra, E.R. & Webster, M.S., 2017. A comparative assessment of SNP and microsatellite markers for assigning parentage in a socially monogamous bird. *Molecular Ecology Resources* 17: 183–193.

Kassis, J.A., Kennison, J.A. & Tamkun, J.W., 2017. Polycomb and trithorax group genes in

Drosophila. *Genetics*, 206(4), pp. 1699–1725.

Khosla, S., Mendiratta, G. & Brahmachari, V., 2006. Genomic imprinting in the mealybugs. *Cytogenetic and genome research*, 113(1–4), pp. 41–52.

Kincaid-Smith, J. *et al.*, 2018. Parent-of-origin-dependent gene expression in male and female schistosome parasites. *Genome Biology and Evolution*, 10(3), pp. 840–856.

Kirkness, E.F. *et al.*, 2010. Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proceedings of the National Academy of Sciences*, 107(27), pp. 12168–12173.

Kitchin, R.M., 1975. Intranuclear destruction of heterochromatin in two species of armored scale insects. *Genetica*, 45(2), pp. 227–235.

Kocher, S.D. *et al.*, 2015. A search for parent-of-origin effects on honey bee gene expression. *G3: Genes, Genomes, Genetics*, 5(8), pp. 1657–1662.

Kol-Maimon, H. *et al.*, 2010. Male behaviors reveal multiple pherotypes within vine mealybug *Planococcus ficus* (Signoret) (Hemiptera; Pseudococcidae) populations. *Naturwissenschaften*, 97(12), pp. 1047–1057.

Kol-Maimon, H., Mendel, Z., *et al.*, 2014a. Paternal inheritance in mealybugs (Hemiptera: Coccoidea: Pseudococcidae). *Naturwissenschaften*, 101(10), pp. 791–802.

Kol-Maimon, H., Ghanim, M., *et al.*, 2014b. Evidence for gene flow between two sympatric mealybug species (Insecta; Coccoidea; Pseudococcidae). *PLoS One*, 9(2), pp. e88433-10.

Kraaijeveld, K., 2009. Male genes with nowhere to hide; sexual conflict in haplodiploids. *Animal Biology*, 59(4), pp. 403–415.

Kuijper, B. & Pen, I., 2010. The evolution of haplodiploidy by male-killing endosymbionts: importance of population structure and endosymbiont mutualisms. *Journal of Evolutionary Biology*, 23(1), pp. 40–52.

Kurtovic, A., Widmer, A. & Dickson, B.J., 2007. A single class of olfactory neurons mediates behavioural responses to a *Drosophila* sex pheromone. *Nature*, 446(7135), p.542.

Kwon, D.H. *et al.*, 2014. Identification and characterization of an esterase involved in malathion resistance in the head louse *Pediculus humanus capitis*. *Pesticide Biochemistry and Physiology*, 112, pp. 13–18.

Lake, C.M. *et al.*, 2007. A genetic analysis of the *Drosophila mcm5* gene defines a domain specifically required for meiotic recombination. *Genetics*, 176(4), pp. 2151–2163.

Larrac*uente, A.M. & Presgraves, D.C., 2012. The selfish Segregation Distorter gene complex of Drosophila melanogaster. Genetics*, 192(1), pp. 33–53.

Leader, D.P. *et al.*, 2017. FlyAtlas 2: a new version of the *Drosophila melanogaster* expression atlas with RNA-Seq, miRNA-Seq and sex-specific data. *Nucleic acids research*, 46, pp. D809–D815.

Leng, N. *et al.*, 2013. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics*, 29(8), pp. 1035–1043.

Lenormand, T. *et al*., 2016. Evolutionary mysteries in meiosis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1706), p.20160001.

Li, B. & Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics*, 12(1), p.323.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Li, H. & Durbin, R., 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, 26(5), pp. 589–595.

Li, H. *et al*., 2015. Higher-level phylogeny of paraneopteran insects inferred from mitochondrial genome sequences. *Scientific Reports*, 5, p. 8527.

Li, W. *et al*., 2010. Genotyping of human lice suggests multiple emergences of body lice from local head louse populations. *PLoS Neglected Tropical Diseases*, 4(3), p.e641.

Light, J.E., Toups, M.A. & Reed, D.L., 2008. What's in a name: The taxonomic status of human head and body lice. *Molecular Phylogenetics and Evolution*, 47(3), pp. 1203–1216.

Lindholm, A.K. *et al*., 2016. The ecology and evolutionary dynamics of meiotic drive. *Trends in Ecology and Evolution*, 31(4), pp. 315–326.

Lo, S.M., Ahuja, N.K. & Francis, N.J., 2009. Polycomb group protein Suppressor 2 of zeste is a functional homolog of Posterior Sex Combs. *Molecular and Cellular Biology*, 29(2), pp. 515–525.

Löfstedt, C., 1993. Moth pheromone genetics and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 340(1292), pp. 167–177.

Lorick, G., 1970. Differential DNA synthesis in heterochromatic and euchromatic chromosome sets of *Planococcus citri*. *Chromosoma*, 32(1), pp. 11–30.

Lott, S.E. *et al*., 2011. Noncanonical compensation of zygotic X transcription in early *Drosophila melanogaster* development revealed through single-embryo RNA-seq. *PLoS Biology*, 9(2), p. e1000590.

Luo, M. *et al*., 2011. A genome-wide survey of imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm. *PLoS Genetics*, 7(6), p.e1002125.

Lyon, M.F., 1961. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature*, 190(4773), p. 372.

Lyon, M.F., 2003. Transmission ratio distortion in mice. *Annual Review of Genetics*, 37(1), pp. 393–408.

Lyttle, T.W., 1993. Cheaters sometimes prosper: distortion of mendelian segregation by meiotic drive. *Trends in Genetics*, 9(6), pp. 205–210.

MacDonald, W.A., 2012. Epigenetic mechanisms of genomic imprinting: common themes in the regulation of imprinted regions in mammals, plants, and insects. *Genetics Research International*, 2012.

Makki, R., Cinnamon, E. & Gould, A.P., 2014. The development and functions of oenocytes. *Annual Review of Entomology*, 59, pp. 405–425.

Malausa, T. *et al.*, 2011. DNA markers to disentangle complexes of cryptic taxa in mealybugs (Hemiptera: Pseudococcidae). *Journal of Applied Entomology*, 135(1-2), pp. 142–155.

Marks, H. *et al.*, 2015. Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. *Genome Biology*, 16(1), p.149.

Martins, R.F. *et al.*, 2012. Isolation and characterization of fifteen polymorphic microsatellite loci for the citrus mealybug, *Planococcus citri* (Hemiptera: Pseudococcidae), and cross-amplification in two other mealybug species. *Journal of Genetics*, 93(1), pp. 75–78.

McKenna, A. *et al.*, 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*.

McManus, C.J. *et al.*, 2010. Regulatory divergence in *Drosophila* revealed by mRNA-seq. *Genome Research*.

McMeniman, C.J. & Barker, S.C., 2005. Transmission ratio distortion in the human body louse, *Pediculus humanus* (Insecta: Phthiraptera). *Heredity*, 96(1), p.63.

Meiklejohn, C.D. & Tao, Y., 2010. Genetic conflict and sex chromosome evolution. *Trends in Ecology and Evolution*, 25(4), pp. 215–223.

Meinking, T.L., 1999. Infestations. *Current Problems in Dermatology*, 11(3), pp. 73–118.

Menon, D.U. & Meller, V., 2010. Germ line imprinting in *Drosophila*: epigenetics in search of function. *Fly*, 4(1), pp. 48–52.

Metsalu, T. *et al.*, 2014. Using RNA sequencing for identifying gene imprinting and random monoallelic expression in human placenta. *Epigenetics*, 9(10), pp. 1397–1409.

Miller, D.R. & Kosztarab, M., 1979. Recent advances in the study of scale insects. *Annual Review of Entomology*, 24(1), pp. 1–27.

Misof, B. *et al.*, 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210), pp. 763–767.

Moore, T. & Haig, D., 1991. Genomic imprinting in mammalian development: a parental tug-of-war. *Trends in Genetics*, 7(2), pp. 45–49.

Myers, E.L., 1932. Two economic greenhouse mealybugs of Mississippi. *Journal of Economic Entomology*, 25(4), pp. 891–896.

Nelson-Rees, W.A., 1962. The effects of radiation damaged heterochromatic chromosomes on male fertility in the mealy bug, *Planococcus citri* (Risso). *Genetics*, 47(6), pp. 661–683.

Nelson-Rees, W.A., Hoy, M.A. & Roush, R.T., 1980. Heterochromatinization, chromatin elimination and haploidization in the parahaploid mite *Metaseiulus occidentalis* (Nesbitt) (Acarina: Phytoseiidae). *Chromosoma*, 77(3), pp. 263–276.

Normark, B.B., 2003. The evolution of alternative genetic systems in insects. *Annual Review of Entomology*, 48(1), pp. 397–423.

Normark, B.B., 2004. Haplodiploidy as an outcome of coevolution between male-killing cytoplasmic elements and their hosts. *Evolution*, 58(4), pp. 790–798.

Normark, B.B., 2006. Perspective: maternal kin groups and the origins of asymmetric genetic systems—genomic imprinting, haplodiploidy, and parthenogenesis. *Evolution*, 60(4), pp. 631–642.

Normark, B.B., 2014. Modes of reproduction. In The evolution of insect mating systems, pp. 1–19. Oxford University Press, USA,

Normark, B.B. & Ross, L., 2014. Genetic conflict, kin and the origins of novel genetic systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1642), p. 20130364.

Norton, R.A. *et al.*, 1993. Phylogenetic perspectives on genetic systems and reproductive modes of mites. In Evolution and diversity of sex ratio in insects and mites., pp. 8–99. Chapman Hall, USA

Nur, U., 1962a. A supernumerary chromosome with an accumulation mechanism in the lecanoid genetic system. *Chromosoma*, 13(3), pp. 249–271.

Nur, U., 1962b. Sperms, sperm bundles and fertilization in a mealy bug, *Pseudococcus obscurus* Essig.(Homoptera: Coccoidea). *Journal of Morphology*, 111(2), pp. 173–199.

Nur, U., 1966. Nonreplication of heterochromatic chromosomes in a mealy bug, *Planococcus citri* (Coccoidea: Homoptera). *Chromosoma*, 19(4), pp. 439–448.

Nur, U., 1967. Reversal of heterochromatization and the activity of the paternal chromosome set in the male mealy bug. *Genetics*, 56(3), pp. 375–389.

Nur, U., 1968. Endomitosis in the mealy bug, *Planococcus citri* (Homoptera: Coccoidea). *Chromosoma*, 24(2), pp. 202–209.

Nur, U., 1970. Translocations between eu-and heterochromatic chromosomes, and spermatocytes lacking a heterochromatic set in male mealy bugs. *Chromosoma*, 29(1), pp. 42–61.

Nur, U., 1972. Origin of mono-and binucleate cells with heterochromatic chromosomes in the malpighian tubules of the european elm scale, *Gossyparia spuria* (Coccoidea: Homoptera). *Journal of Morphology*, 137(2), pp. 243–256.

Nur, U., 1980. Evolution of unusual chromosome systems in scale insects (Coccoidea: Homoptera). In Insect cytogenetics, pp. 97–118.. Blackwell, United Kingdom.

Nur, U., 1990. Heterochromatization and euchromatization of whole genomes in scale insects (Coccoidea: Homoptera). *Development*, 108, pp. 29–34.

Nur, U. & Chandra, H.S., 1963. Interspecific hybridization and gynogenesis in mealy bugs. *The American Naturalist*, 97(895), pp. 197–202.

O'Brien, T.W., 2002. Evolution of a protein-rich mitochondrial ribosome: implications for human genetic disease. *Gene*, 286(1), pp. 73–79.

O'Brien, T.W., 2003. Properties of human mitochondrial ribosomes. *IUBMB Life*, 55(9), pp. 505–

513.

Oakeshott, J.G. *et al.*, 2003. The genomics of insecticide resistance. *Genome Biology*, 4(1), p.202.

Olds, B.P. *et al.*, 2012. Comparison of the transcriptional profiles of head and body lice. *Insect Molecular Biology*, 21(2), pp. 257–268.

Orr, H.A., 1996. Dobzhansky, Bateson, and the genetics of speciation. *Genetics*, 144(4), pp. 1331–1335.

Otto, S.P., 2009. The evolutionary enigma of sex. *The American Naturalist*, 174(S1), pp. S1–S14.

Page Jr, R.E. & Amdam, G. V, 2007. The making of a social insect: developmental architectures of social design. *Bioessays*, 29(4), pp. 334–343.

Palazzo, A.F. & Lee, E.S., 2015. Non-coding RNA: what is functional and what is junk? *Frontiers in Genetics*, 6, p.2.

Park, J.W. *et al.*, 2004. Identification of alternative splicing regulators by RNA interference in *Drosophila*. *Proceedings of the National Academy of Sciences*, 101(45), pp. 15974–15979.

Parker, J.D. & Hedrick, P.W., 2000. Gene flow and selection balance in haplodiploid social insects. *Heredity*, 85(6), p.530.

Patten, M.M. *et al.*, 2014. The evolution of genomic imprinting: theories, predictions and empirical tests. *Heredity*, 113(2), p.119.

Perotti, M.A. *et al.*, 2004. The sex ratio distortion in the human head louse is conserved over time. *BMC Genetics*, 5(1), p.1.

Phadnis, N. & Orr, H.A., 2009. A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids. *Science*, 323(5912), pp. 376–379.

Plath, K. *et al.*, 2003. Role of histone H3 lysine 27 methylation in X inactivation. *Science*, 300(5616), pp. 131–135.

Pollard, K.S. *et al.*, 2008. A genome-wide approach to identifying novel-imprinted genes. *Human Genetics*, 122(6), pp. 625–634.

Pompanon, F. *et al.*, 2005. Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, 6(11), p.847.

Prantera, G. & Bongiorni, S., 2012. Mealybug chromosome cycle as a paradigm of epigenetics. *Genetics Research International*, 2012.

Previte, D. *et al.*, 2014. Differential gene expression in laboratory strains of human head and body lice when challenged with *Bartonella quintana*, a pathogenic bacterium. *Insect Molecular Biology*, 23(2), pp. 244–254.

Queller, D.C., 2003. Theory of genomic imprinting conflict in social insects. *BMC Evolutionary Biology*, 3(1), p. 15.

Quinlan, A.R. & Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic

features. *Bioinformatics*, 26(6), pp. 841–842.

Rackham, O. & Filipovska, A., 2014. Supernumerary proteins of mitochondrial ribosomes. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1840(4), pp. 1227–1232.

Raju, N.B., 1994. Ascomycete spore killers: chromosomal elements that distort genetic ratios among the products of meiosis. *Mycologia*, pp. 461–473.

Raoult, D. & Roux, V., 1999. The body louse as a vector of reemerging human diseases. *Clinical Infectious Diseases*, 29(4), pp. 888–911.

Raymond, M. & Rousset, F., 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86(3), pp. 248–249.

Reik, W. & Walter, J., 2001. Genomic imprinting: parental influence on the genome. *Nature Reviews Genetics*, 2(1), p.21.

Remnant, E.J. *et al.*, 2016. Parent-of-origin effects on genome-wide DNA methylation in the Cape honey bee (*Apis mellifera capensis*) may be confounded by allele-specific methylation. *BMC Genomics*, 17(1), p.226.

Rice, W.R., 2013. Nothing in genetics makes sense except in light of genomic conflict. *Annual Review of Ecology, Evolution, and Systematics*, 44, pp. 217–237.

Richman, T.R., Rackham, O. & Filipovska, A., 2014. Mitochondria: unusual features of the mammalian mitoribosome. *The International Journal of Biochemistry & Cell biology*, 53, pp. 115–120.

Roelofs, W. *et al.*, 1987. Sex pheromone production and perception in European corn borer moths is determined by both autosomal and sex-linked genes. *Proceedings of the National Academy of Sciences*, 84(21), pp. 7585–7589.

Ross, L., Pen, I. & Shuker, D.M., 2010a. Genomic conflict in scale insects: the causes and consequences of bizarre genetic systems. *Biological Reviews*, 85(4), pp. 807–828.

Ross, L., Langenhof, M.B.W., *et al.*, 2010b. Sex allocation in a species with paternal genome elimination: the roles of crowding and female age in the mealybug *Planococcus citri*. *Evolutionary Ecology Research*, 12(1), pp. 89–104.

Ross, L. *et al.*, 2012. The role of endosymbionts in the evolution of haploid-male genetic systems in scale insects (Coccoidea). *Ecology and Evolution*, 2(5), pp. 1071–1081.

Ross, L. *et al.*, 2013. Large population size predicts the distribution of asexuality in scale insects. *Evolution*, 67(1), pp. 196–206.

Ross, L. & Normark, B.B., 2015. Evolutionary problems in centrosome and centriole biology. *Journal of Evolutionary Biology*, 28(5), pp. 995–1004.

Rotundo, G. & Tremblay, E., 1982. Hybridization and sex pheromone responses in two closely related mealybug species (Homoptera: Pseudococcidae). *Systematic Entomology*, 7(4), pp. 475–478.

Rousset, F., 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, 8(1), pp. 103–106.

Sabour, M., 1972. RNA synthesis and heterochromatization in early development of a mealybug. *Genetics*, 70(2), pp. 291–298.

Santoni, F.A. *et al.*, 2017. Detection of imprinted genes by single-cell allele-specific gene expression. *The American Journal of Human Genetics*, 100(3), pp. 444–453.

Sasaki-Fukatsu, K. *et al.*, 2006. Symbiotic bacteria associated with stomach discs of human lice. *Applied and Environmental Microbiology*, 72(11), pp. 7349–7352.

Schimenti, J., 2000. Segregation distortion of mouse t haplotypes: the molecular basis emerges. *Trends in Genetics*, 16(6), pp. 240–243.

Scholl, H., 1955. Ein Beitrag zur Kenntnis der Spermatogenese der Mallophagen. *Chromosoma*, 7(1), pp. 271–274.

Schrader, F., 1921. The chromosomes of *Pseudococcus nipae*. *The Biological Bulletin*, 40(5), pp. 259–269.

Schrader, F., 1935. Notes an the mitotic behavior of long chromosomes. *Cytologia*, 6(4), pp. 422–430.

Schrader, F. & Hughes-Schrader, S., 1931. Haploidy in metazoa. *The Quarterly Review of Biology*, 6(4), pp. 411–438.

Schulten, G.G.M., 1985. Pseudo-arrhenotoky. In Spider mites, their biology, natural enemies and control, world crop pests pp. 67–71.

Schurko, A.M., Neiman, M. & Logsdon Jr, J.M., 2009. Signs of sex: what we know and how we know it. *Trends in Ecology & Evolution*, 24(4), pp. 208–217.

Schwander, T. & Oldroyd, B.P., 2016. Androgenesis: where males hijack eggs to clone themselves. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1706), p. 20150534.

Schwartz, Y.B. & Pirrotta, V., 2008. Polycomb complexes and epigenetic states. *Current Opinion in Cell Biology*, 20(3), pp. 266–273.

Shao, R., Kirkness, E.F. & Barker, S.C., 2009. The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, *Pediculus humanus*. *Genome Research*.

Shin, S. *et al.*, 2013. Molecular phylogeny of black fungus gnats (Diptera: Sciaroidea: Sciaridae) and the evolution of larval habitats. *Molecular Phylogenetics and Evolution*, 66(3), pp. 833–846.

Shuker, D.M., Moynihan, A.M. & Ross, L., 2009. Sexual conflict, sex allocation and the genetic system. *Biology Letters*, 5(5), pp. 682–685.

Simão, F.A. *et al.*, 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), pp. 3210–3212.

Smith, S., Witkowski, A. & Joshi, A.K., 2003. Structural and functional organization of the animal fatty acid synthase. *Progress in Lipid Research*, 42(4), pp. 289–317.

Stanley-Samuelson, D.W. *et al.*, 1988. Fatty acids in insects: composition, metabolism, and biological significance. *Archives of Insect Biochemistry and Physiology*, 9(1), pp. 1–33.

Stouthamer, R. *et al.*, 2010. Intragenomic conflict in populations infected by parthenogenesis inducing *Wolbachia* ends with irreversible loss of sexual reproduction. *BMC Evolutionary Biology*, 10(1), p.229.

Takano-Lee, M. *et al.*, 2003. In vivo and in vitro rearing of *Pediculus humanus capitis* (Anoplura: Pediculidae). *Journal of Medical Entomology*, 40(5), pp. 628–635.

Takano-Lee, M. *et al.*, 2005. Transmission potential of the human head louse, *Pediculus capitis* (Anoplura: Pediculidae). *International Journal of Dermatology*, 44(10), pp. 811–816.

Tao, Y. *et al.*, 2007. A sex-ratio meiotic drive system in *Drosophila simulans*. I: an autosomal suppressor. *PLoS Biology*, 5(11), p.e292.

Timmons, J.A., Szkop, K.J. & Gallagher, I.J., 2015. Multiple sources of bias confound functional enrichment analysis of global-omics data. *Genome Biology*, 16(1), p.186.

Tombesi, M.L. & Papeschi, A.G., 1993. Meiosis in Haematopinus suis and Menacanthus stramineus (Phthiraptera, Insecta). *Hereditas*, 119(1), pp. 31–38.

Tombesi, M.L., Papeschi, A.G. & Mola, L.M., 1999. Spermatogenesis in *Bovicola limbata* Gervais, 1844 and *B. caprae* Gurlt, 1843 (Phthiraptera, Ischnocera). *Cytologia*, 64(1), pp. 25–27.

Tranfaglia, A. & Tremblay, E., 1982. A morphological comparison between *Planococcus citri* (Ris.), *Planococcus ficus* (Sign.) and their F1 hybrids (Homoptera: Coccoidea: Pseudococcidae). *Entomotaxonomia*, 4(1/2), pp. 1–5.

Trivers, R.L. & Hare, H., 1976. Haploidploidy and the evolution of the social insect. *Science*, 191(4224), pp. 249–263.

Tselykh, T. V, Roos, C. & Heino, T.I., 2005. The mitochondrial ribosome-specific MrpL55 protein is essential in *Drosophila* and dynamically required during development. *Experimental Cell Research*, 307(2), pp. 354–366.

Turelli, M. & Moyle, L.C., 2007. Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics*.

Úbeda, F. & Wilkins, J.F., 2008. Imprinted genes and human disease: an evolutionary perspective. In Genomic imprinting, pp. 101–115. Springer, USA.

Úbeda, F., Patten, M.M. & Wild, G., 2015. On the origin of sex chromosomes from meiotic drive. *Proceedings of the Royal Society of London B: Biological Sciences*, 282(1798), p.20141932.

Veracx, A. & Raoult, D., 2012. Biology and genetics of human head and body lice. *Trends in Parasitology*, 28(12), pp. 563–571.

Viera, A., Page, J. & Rufas, J., 2008. Inverted meiosis: the true bugs as a model to study. *Meiosis*, 5, pp. 137–156.

Volpi, S., Bongiorni, S. & Prantera, G., 2007. HP2-like protein: a new piece of the facultative heterochromatin puzzle. *Chromosoma*, 116(3), pp. 249–258.

Von Dohlen, C.D. *et al.*, 2001. Mealybug β-proteobacterial endosymbionts contain γ-proteobacterial symbionts. *Nature*, 412(6845), p.433.

Wade, C.M. & Daly, M.J., 2005. Genetic variation in laboratory mice. *Nature Genetics*, 37(11), p.1175.

Wang, X. *et al.*, 2008. Transcriptome-wide identification of novel imprinted genes in neonatal mouse brain. *PLoS One*, 3(12), p.e3839.

Wang, X., Soloway, P.D. & Clark, A.G., 2011. A Survey for Novel Imprinted Genes in the Mouse Placenta by mRNA-seq. *Genetics*, 189(1), pp. 109–122.

Wang, X. *et al.*, 2012. Random X inactivation in the mule and horse placenta. *Genome Research*, p. gr-138487.

Wang, X. & Clark, A.G., 2014. Using next-generation RNA sequencing to identify imprinted genes. *Heredity*, 113(2), pp. 156–166. Available at: http://dx.doi.org/10.1038/hdy.2014.18.

Wang, X., Werren, J.H. & Clark, A.G., 2016. Allele-specific transcriptome and methylome analysis reveals stable inheritance and *cis*-regulation of DNA methylation in *Nasonia*. *PLoS Biology*, 14(7), pp. e1002500-21.

Waters, A.J. *et al.*, 2011. Parent-of-origin effects on gene expression and DNA methylation in the maize endosperm. *The Plant Cell*, 23(12), pp. 4221–4233.

Weir, B.S. & Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, pp. 1358–1370.

Werren, J.H., 1993. The evolution of inbreeding in haplodiploid organisms. In The natural history of inbreeding and outbreeding, pp. 42–59.

Werren, J.H., 2011. Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proceedings of the National Academy of Sciences*, 108(Supplement 2), pp. 10863–10870.

Werren, J.H. & Beukeboom, L.W., 1998. Sex determination, sex ratios, and genetic conflict. *Annual Review of Ecology and Systematics*, 29(1), pp. 233–261.

Werren, J.H., Baldo, L. & Clark, M.E., 2008. *Wolbachia*: master manipulators of invertebrate biology. *Nature Reviews Microbiology*, 6(10), p.741.

Wilkins, J.F., Úbeda, F. & Van Cleve, J., 2016. The evolving landscape of imprinted genes in humans and mice: conflict among alleles, genes, tissues, and kin. *BioEssays*, 38(5), pp. 482–489.

Wittkopp, P.J. & Kalay, G., 2012. *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics*, 13(1), p.59.

Wrensch, D.L., Kethley, J.B. & Norton, R.A., 1994. Cytogenetics of holokinetic chromosomes and inverted meiosis: keys to the evolutionary success of mites, with generalizations on eukaryotes. *Mites*, pp. 282–343. Springer, USA.

Wright, S., 1931. Evolution in Mendelian populations. *Genetics*, 16(2), pp. 97–159.

Wright, S., 1933. Inbreeding and homozygosis. *Proceedings of the National Academy of*

Sciences, 19(4), pp. 411–420.

Wutz, A., 2011. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nature Reviews Genetics*, 12(8), p.542.

Yang, F. *et al.*, 2010. Global survey of escape from X inactivation by RNA-sequencing in mouse. *Genome Research*, 20(5), pp. 614–622.

Yokogawa, T. & Yahara, T., 2009. Mitochondrial phylogeny certified PGL (Paternal Genome Loss) is of single origin and haplodiploidy sensu stricto (arrhenotoky) did not evolve from PGL in the scale insects (Hemiptera: Coccoidea). *Genes & Genetic Systems*, 84(1), pp. 57–66.

Yoon, K.S. *et al.*, 2006. An improved in vitro rearing system for the human head louse allows the determination of resistance to formulated pediculicides. *Pesticide Biochemistry and Physiology*, 86(3), pp. 195–202.

Yoon, K.S. *et al.*, 2008. Three mutations identified in the voltage-sensitive sodium channel α-subunit gene of permethrin-resistant human head lice reduce the permethrin sensitivity of house fly Vssc1 sodium channels expressed in *Xenopus* oocytes. *Insect Biochemistry and Molecular Biology*, 38(3), pp. 296–306.

Yoshizawa, K. & Johnson, K.P., 2003. Phylogenetic position of Phthiraptera (Insecta: Paraneoptera) and elevated rate of evolution in mitochondrial 12S and 16S rDNA. *Molecular Phylogenetics and Evolution*, 29(1), pp. 102–114.

Yoshizawa, K. & Johnson, K.P., 2010. How stable is the “Polyphyly of Lice” hypothesis (Insecta: Psocodea)?: A comparison of phylogenetic signal in multiple genes. *Molecular Phylogenetics and Evolution*, 55(3), pp. 939–951.

Zhuo, Z., Lamont, S.J. & Abasht, B., 2017. RNA-seq analyses identify frequent allele specific expression and no evidence of genomic imprinting in specific embryonic tissues of chicken. *Scientific Reports*, 7(1), p. 11944.

Appendix 1

Haplodiploidy and the reproductive ecology of arthropods

As published:

De la Filia, A.G., Bain, S.A. & Ross, L., 2015.

Haplodiploidy and the reproductive ecology of Arthropods.

Current Opinion in Insect Science, 9, pp. 36–43.

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Insect Science

Haplodiploidy and the reproductive ecology of Arthropods

Andrés G de la Filia[1], Stevie A Bain[1] and Laura Ross

Approximately 15% of all arthropods reproduce through haplodiploidy. Yet it is unclear how this mode of reproduction affects other aspects of reproductive ecology. In this review we outline predictions on how haplodiploidy might affect mating system evolution, the evolution of traits under sexual or sexual antagonistic selection, sex allocation decisions and the evolution of parental care. We also give an overview of the phylogenetic distribution of haplodiploidy. Finally, we discuss how comparisons between different types of haplodiploidy (arrhenotoky, PGE with haploid vs *somatically* diploid males) might help to discriminate between the effects of virgin birth, haploid gene expression and those of haploid gene transmission.

Address
School of Biological Sciences, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, UK

Corresponding author: Ross, Laura (laura.ross@ed.ac.uk)

[1] Authors contributed equally.

Current Opinion in Insect Science 2015, **9**:36–43

This review comes from a themed issue on **Behavioural ecology**

Edited by **Cécile Lelann** and **Jacintha Ellers**

doi:10.1016/j.cois.2015.04.018

Introduction

Behavioural ecology focuses on understanding how natural selection shapes the way organisms behave. Insects have featured prominently as model systems [1]. Despite providing important general insights, these studies fail to include the full diversity of reproductive systems in arthropods. Most assume that each parent is contributing an equal share of their genes to their offspring. Yet as many as 15% of arthropods are haplodiploids [2–4], where mothers monopolize the production of male offspring, either by the asexual production of sons (arrhenotoky) or by producing sons that eliminate their father's genome from their germline (paternal genome elimination, PGE) [4]. Haplodiploidy has received attention in the context of eusociality (though its importance has increasingly fallen out of favour [5]), yet how it affects other aspects of species' ecology has barely been addressed. Here we consider its role in reproductive behaviour and mating system evolution. We summarize available theory (main text and Table 1) and empirical data (supplementary Table S1), provide verbal models when formal ones are lacking, and identify areas that need addressing in the future.

Most biologists are familiar with haplodiploidy in the Hymenoptera. The vast majority of hymenopterans reproduce through arrhenotoky [4], and most behavioural ecology studies on haplodiploid species involve members of this order. Yet, it constitutes just one of two-dozen independent origins of haplodiploidy [4]: arrhenotoky is also found among thrips, some hemipterans and several clades of beetles and mites. PGE, where males develop from fertilized eggs but subsequently eliminate the paternal chromosomes, is found in most scale insects (Hemiptera), some beetles, flies, springtails, lice and mites (in total about 20 000 species) [6] (Figures 1 and 2). Different PGE species vary in the timing of the elimination of the paternal genome, and in whether it becomes transcriptionally silenced or not [4,6]. As a result, male gene expression varies from haploid to diploid (Figure 3) with various intermediates. This variability is important as it might allow differentiation of the effects of haploid gene transmission and those of haploid gene expression, while comparisons between arrhenotokous and PGE taxa could provide insights into the importance of virgin birth (see Table 1).

Evolutionary genetics under haplodiploidy

Haplodiploidy affects the evolutionary genetics of species in a number of ways. Under arrhenotoky and some types of PGE, gene expression in males is haploid and maternal. Therefore, recessive mutations are exposed to selection in males, firstly, reducing genetic load, due to a lower effective mutation rate and the exposure of deleterious recessive alleles in haploid males [7] and secondly, increasing the rate at which rare recessive beneficial mutations can spread. As a result, these species are expected to adapt faster to changing environments. This is true only for non-sex specific traits. The evolution of male-limited traits is complex, as sons do not inherit them from their fathers (Figure 3). In addition, selection among females has a relatively greater impact on evolutionary change as each gene finds itself more frequently in females than males [8].

Sexual selection

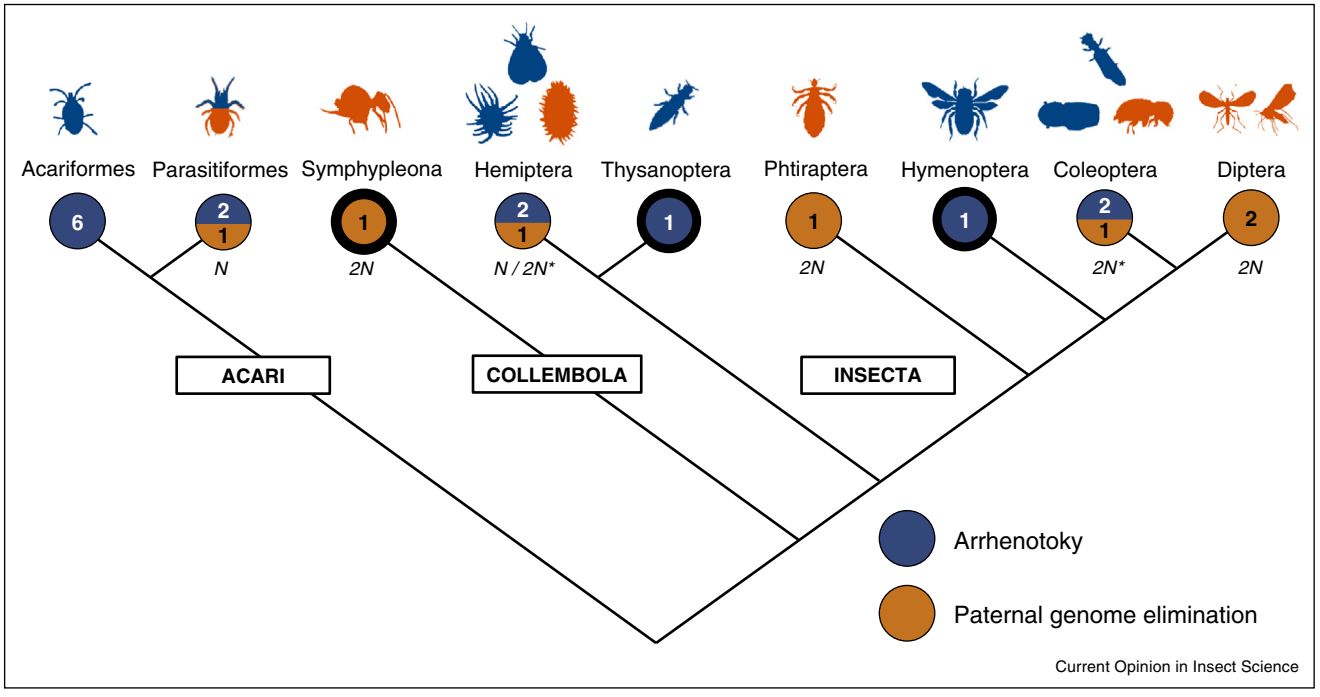
Sexual selection arises through competition within a sex (usually males) for access to mates (and their gametes) [9]

Table 1

An overview of the prediction for each of the reproductive ecology traits discussed in the manuscript. We summarize how we expect the three different types of haplodiploid organisms to differ with respect to diplodiploid taxa. + indicates that trait is promoted relative to diplodiploidy, − that the type of haplodiploidy inhibits the evolution of the trait, while = indicates that there is no expected difference between haplodiploids and diplodiploids. Please note that most of these predictions, especially differences between the different types of haplodiploids, are based on verbal models and will need to be corroborated by formal theory in the future.

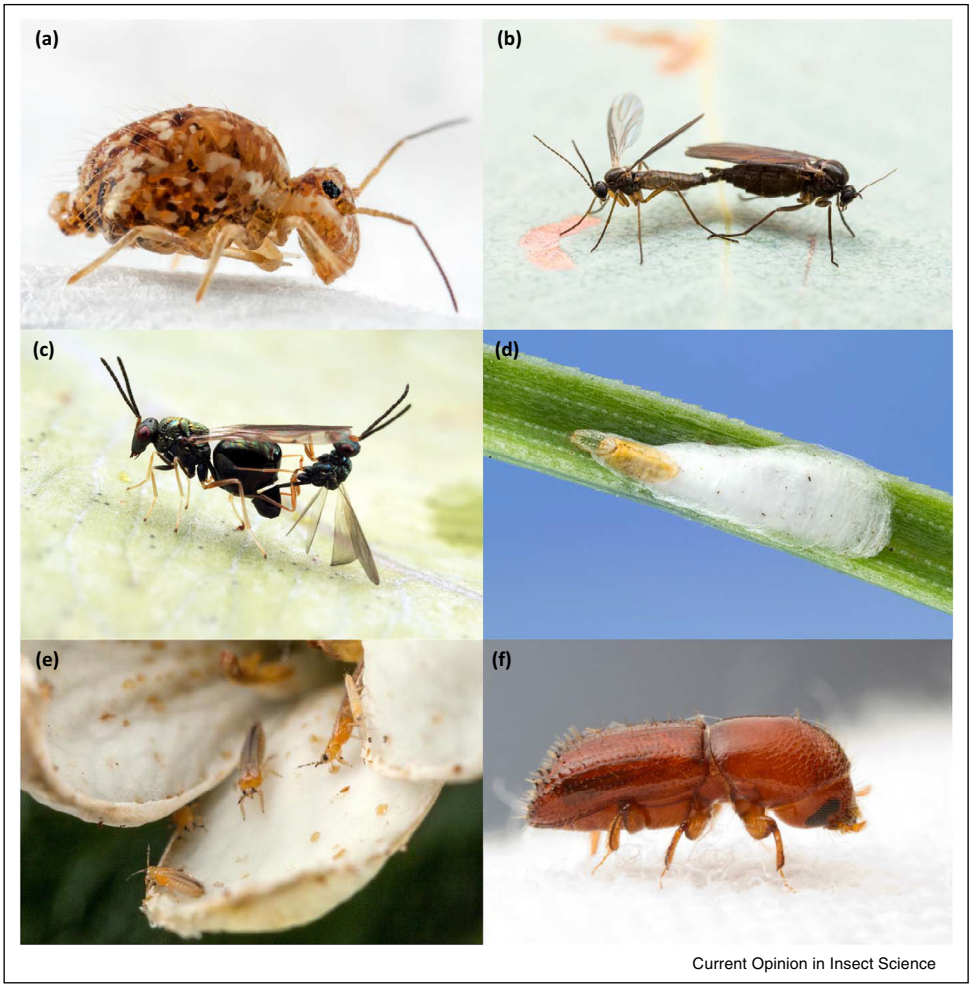
	Arrhenotoky	PGE (haploid soma)	PGE (diploid soma)
Inbreeding depression	− [24]	−	=
Exaggerated sexual selected traits (under Fisherian runaway selection)	− [12]	−	−/= Sons might express their fathers' ornament, increasing their reproductive success, yet are unable to pass it on to their offspring
Exaggerated sexual selected traits (handicap principle)	+ [13]	+	+
Intra-locus conflicts	Resolved in favour of females (dominant traits), polymorphism (recessive traits) [16]	Resolved in favour of females (dominant traits), polymorphism (recessive traits)	Resolved in favour of females (both dominant and recessive traits).
Inter-locus conflicts	Females are more likely to evolve resistance to male trait [16]	Females are more likely to evolve resistance to male trait	= (?) Mothers equally likely to evolve resistance as sons might express their fathers' trait, benefitting the mother through their increased reproductive success
Sperm cooperation	+ [52]	+ [52]	+ [52]
Fertility assurance	+	− Virgin females unable to produce offspring	− Virgin females unable to produce offspring
Facultative sex ratio control	+ [33,53]	+/= [54,55]	+/= [56••]
Polyandry	−/= [11••]	+	+
Maternal care	= [49]	=	=
Paternal care	+ [50••]	+	+

Figure 1



Schematic cladogram of arrhenotokous (blue) and PGE (orange) groups in Arthropoda. The number of independent origins of haplodiploidy is indicated within the circles. Clades in which all members are haplodiploid are indicated with a black ring around the circle. The type of PGE is indicated below the circle with *2N* for germline PGE, *2N** for germline PGE, where the paternal genome is transcriptionally silenced in somatic cells and *N* for embryonic PGE. Origins outside the Arthropoda (rotifers and nematodes) are not shown.

Figure 2

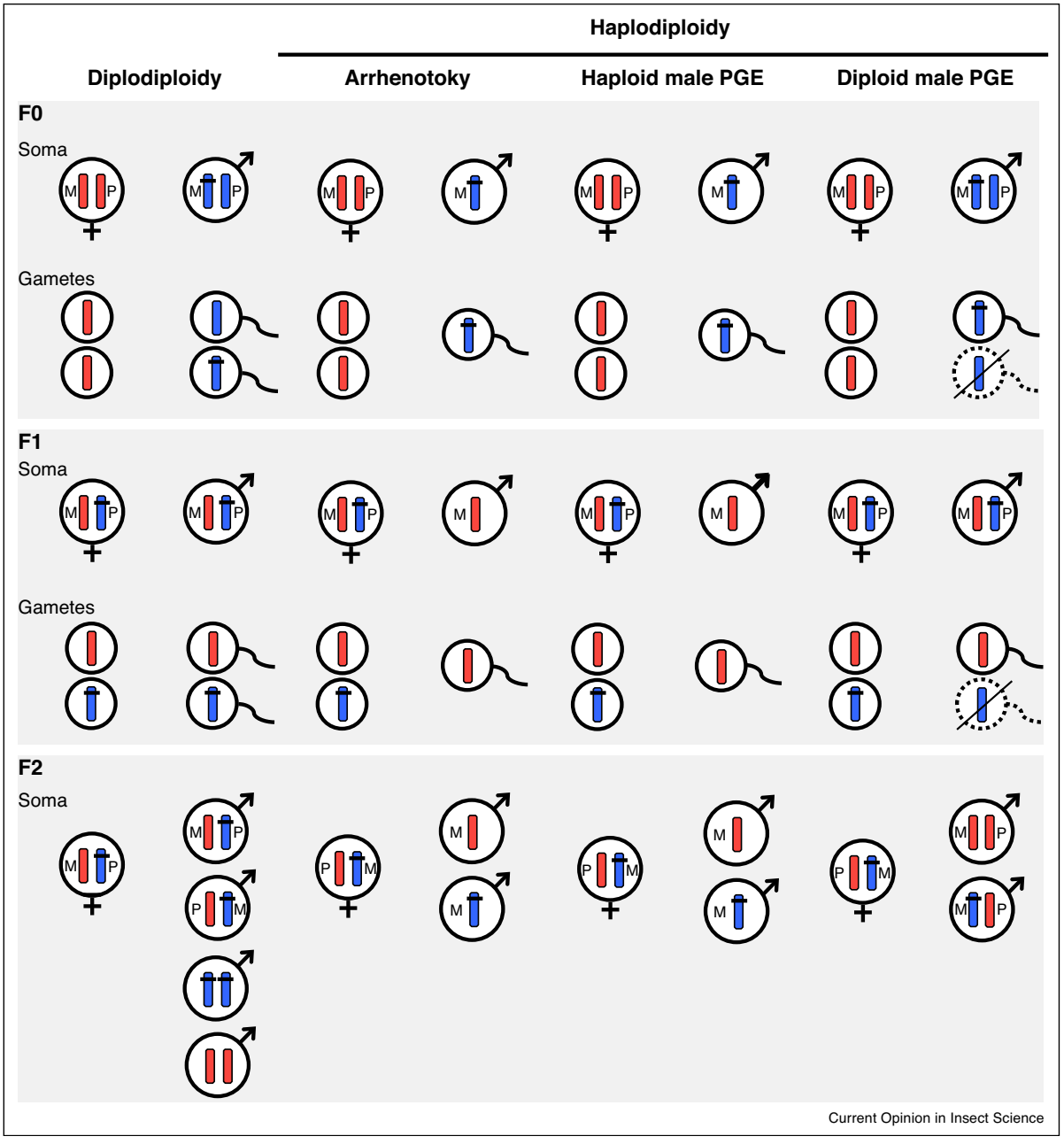


A number of examples of PGE species: **(a)** a globular springtail (Symphypleona), **(b)** a pair of mating fungus gnats (Sciaridae), **(d)** the armoured scale insect *Chionaspis pinifoliae.* And a number of arrhenotokous species: **(c)** Eucharitid wasps mating, **(e)** flower thrips, **(f)** *Xyleborus* sp. ambrosia beetle.
Source: Images b-f © Alex Wild and image a © Gil Wizen, used with permission.

and can result in the evolution of exaggerated traits. Such traits evolve if females chose to mate with males carrying them, either because the trait signals genetic quality (the handicap principle) [10], or because their sons will inherit it and therefore be attractive to other females (Fisherian runaway selection) [9].

As haplodiploid sons do not inherit traits from their fathers their maternal grandfathers are their closest male progenitors, so that selection on male traits skips generations (Figure 3) [11••]. A simulation study [12] showed that, due to this delay, rare alleles encoding male ornaments are particularly likely to be lost through genetic drift. The same might be true for alleles underlying traits that increase a male's reproductive success without being a direct target of female choice, such as combat ability. Subsequent deterministic models showed that haplodiploid transmission genetics also affects the genetic correlation between male traits and female preference, thereby promoting sexual selection through the handicap principle, but impeding Fisherian runaway selection [13]. Together, these models suggest that haplodiploidy should affect the evolution of exaggerated male traits. Comparative efforts to identify the prevalence of such traits and the degree of sexual dimorphism between haplodiploid and diplodiploid species might therefore, in principle, provide insight into the relative importance of runaway versus handicap selection. However these predictions are based on a number of simplifying assumptions, and there is an urgent need for more formal theory considering finite population sizes, costs of female preference, sex-specific mutation rates and allelic dominance.

Figure 3



Genetic inheritance of a paternal allele under diplodiploidy and the three different types of haplodiploidy. Somatic genotypes are represented for three generations (F0, F1 and F2) and gamete genotypes for F0 and F1. (For simplicity, assume that there is no meiotic recombination and that offspring mate to produce the next generation.) Maternal chromosomes in F0 are shown in pink and paternal chromosomes in F0 are shown in blue. Black lines in the upper half of some chromosomes represent a given male trait (e.g., an advantageous trait in inter-locus conflict or sexual selection), whose inheritance we follow across three generations. M, inherited from the mother P, inherited from the father. Colours refer to maternal (red) and paternal (blue) genomes in F0. In F1 and F2, only one out of four possible female soma genotypes are shown, while all possible male soma genotypes carrying the original paternal allele are indicated for both generations. The figure shows how transmission of the male trait is affected by the different genetic systems. Paternal line inheritance is possible under diplodiploidy only. Due to the fact that arrhenotokous males develop from unfertilized eggs, they do not inherit paternal chromosomes. In haploid male PGE, the situation is similar because paternal chromosomes are eliminated in the early developmental stages of the zygote. Also, in certain groups with diploid male PGE, such as in Neococcidae, the paternal genome is transcriptionally repressed and F1 males will not express the male trait. In these three scenarios, selection does not act upon this trait in F1 males.

Sexual conflict

Sexual conflicts result from the discordance of genetic interests between the sexes; males and females might differ in their optimal allele state or expression level at the same gene (intra-locus conflict) [14] or different loci (inter-locus conflict) [15••]. Specific theory on the role of haplodiploidy is scarce, but analogies to X-chromosome inheritance allow us to utilize theory of sexual conflict under sex linkage [16]. As haplodiploid males obtain reproductive success only through daughters, male-beneficial traits that reduce female fitness are particularly unlikely to spread [16]. Intra-locus conflicts, in particular, will tend to be resolved in favour of females. Only if the trait is recessive and its effects are masked in females, could a polymorphism with a female advantageous allele arise [14], though not under PGE with somatically diploid males. By contrast, inter-locus conflict, especially over sex-limited traits, will not necessarily be resolved in favour of females. For example, a trait that reduces female fecundity but increases sperm competitive ability is as likely to spread under haplodiploidy as under diplodiploidy [15••], yet, while diplodiploid females could benefit indirectly, through sons inheriting it, haplodiploid mothers cannot. As a result haplodiploid females are more likely to evolve resistance mechanisms [16]. Unfortunately there are few empirical studies on either intra-locus or inter-locus sexual conflicts and the predictions outlined here remain to be corroborated.

Sperm cooperation

In diploids, each individual sperm carries a unique haploid genome, different from the diploid genome of the male [17, 18••,19]. Under haplodiploidy sperm are produced mitotically, so individual sperm are genetically identical (barring mutations). As a result, there might be more scope for sperm cooperation, especially under post-copulatory sexual selection [18••]. Empirical data on sperm behaviour under haplodiploidy are limited. However, a peculiar type of sperm cooperation has been found among scale insects with PGE. In this group, individual sperm cells have lost their motility, which they regain by assembling into motile sperm bundles, consisting of tens or even hundreds of sperm cells [20].

Mating systems and inbreeding

There is a strong empirical association between the occurrence of haplodiploidy and certain mating systems, especially those in which inbreeding is systematic (Table S1). Examples include arrhenotokous species such as many parasitoid wasps or bark and ambrosia beetles [21,22], and PGE species such as the coffee-borer beetle [23]. Haplodiploids are more resistant to inbreeding depression due to their reduced genetic load [24,25••,26,27]. However, inbreeding can be detrimental under some conditions: hymenopterans with complementary sex determination (CSD) are greatly affected, as inbreeding produces sterile diploid homozygous males [28]. In PGE species in which the paternal genome is transcriptionally active [29] (diploid male PGE, Figure 3) deleterious recessive alleles are not exposed to selection, so such species are expected to suffer from substantial inbreeding depression. Some of them might have evolved monogeny (where all offspring of each individual female are either exclusively male or exclusively female) as an elaborate mechanisms to avoid inbreeding [30]. Finally we expect substantial inbreeding depression in females under all types of haplodiploidy when inbreeding depression is caused by genes with female-limited expression [27].

Another aspect that could have strong implications on mating systems of arrhenotokous, but not PGE, species is that unmated females can still reproduce by producing all-son broods, which could result in relaxed selection for mate-finding traits compared to diploid/PGE females [31], or allow females to be more choosy. The capacity for virgin birth might also make arrhenotokous females good colonizers: a single arrhenotokous female could theoretically establish a population by producing sons and mating with them. Sex ratio control under arrhenotoky allows for the female-biased sex ratios favoured under such conditions [32,33]. Empirical support comes from ambrosia beetles, where incestuous arrhenotokous species are predominant over diploid outbreeding species with similar ecology on remote islands [34].

Finally, haplodiploidy might affect female mating rates. Monogamy has received considerable attention in the Hymenoptera as an important pre-requisite for the evolution of eusociality. Yet, although there is a huge literature on the link between haplodiploidy and eusociality, few authors have discussed whether monogamy is more or less prevalent among haplodiploids (although see [11••,31,35]). Females are thought to mate multiply to obtain either direct (nuptial gifts, replenishment of sperm supplies) or indirect benefits (promote genetic diversity, increase probability of genetic compatibility) [11••]. In theory, haplodiploidy could affect both. Arrhenotokous females use sperm only to fertilize their female eggs and are able to produce sons without sperm. As a result they might both be less likely to become sperm depleted, and to suffer low reproductive success [22,31,36]. In addition, because the cost of remaining unmated is less severe, females can afford to be choosier about whom to mate with. This is not expected under PGE, as females require sperm to fertilize zygotes of both sexes. In terms of indirect genetic effects, both PGE and arrhenotokous females produce broods that are less genetically diverse on average than diplodiploid females do. They might, therefore, be selected to compensate for this by multiple mating. This is supported by various studies on haplodiploid obligately eusocial species [37,38]. To summarize, female remating rates might vary substantially among haplodiploids, but the relative balance between direct and indirect benefits suggests that, on average, they

would be lowest for arrhenotokous female and highest for those with PGE.

Sex allocation

Sex allocation is perhaps the only aspect of insect reproductive behaviour where studies on haplodiploid species are over-represented. The ability of haplodiploids to precisely alter the sex ratio of their offspring is well documented [39]. Increased control over sex allocation is obvious in haplodiploid taxa, where, unlike under genetic sex determination in diploids, there is no default sex ratio of 50:50. Sex ratio control might have allowed haplodiploid species to evolve a wide range of mating systems and promote alloparental brood care, as mothers are able to bias their sex ratio towards the more helpful sex [40]. However, this flexibility might come at the cost of increased conflicts over sex allocation: First of all, sexual conflict arises between parents over the sex ratio of their offspring [41,42]. Haplodiploid mothers generally favour an equal investment into each sex [43]. Yet fathers, who are not related to male offspring, favour a strongly female biased sex ratio and may evolve ways to persuade their partner to increase fertilization rates (under arrhenotoky) or manipulate the sex determining mechanism (under PGE). Support for the possibility that arrhenotokous males can, under some conditions, manipulate sex allocation decisions of their partners comes from parasitoid wasps [44,45], and spider mites [46••]. Although no studies have yet considered male influence on sex allocation under PGE, it might be more likely to occur as fathers' genes are present in sons [41].

Haplodiploidy might also lead to conflicts among siblings over sex allocation, and between parents and offspring, in those species where siblings interact. Under haplodiploidy, a female is more closely related to her sisters than to her brothers, and should favour a more female-biased sex ratio. The occurrence of these conflicts and how they are resolved has been studied extensively in the eusocial Hymenoptera [47] but have received less attention in other taxa where they are expected to occur, such as social thrips and mites.

Parental care

Another aspect of insect reproduction that varies dramatically between species is the presence of parental care and the relative energy expenditure males and females devote to caring for their young. A number of studies have investigated how haplodiploidy affects the evolution of paternal versus maternal care. A population genetic model by Wade [48] suggested that haplodiploidy facilitates the evolution of maternal care, but assumed that the cost of maternal care rests on both parents, not just on the mother. A subsequent model [49] included the latter possibility as well as effects of inbreeding and alternative assumptions about the genetic underpinning of the maternal care. The result of this model suggests that haplodiploidy does not generally promote maternal care.

What about paternal care? Intuitively, haplodiploidy might be expected to inhibit paternal care as males are selected to care only for their female offspring. However, although fathers value their sons less than under diplodiploidy, they value their daughters more, and these two effects exactly cancel [50••]. So under outbreeding, haplodiploidy neither promotes nor inhibits paternal care. By contrast, under inbreeding haplodiploidy may promote paternal care, as it inflates a male's relatedness to his offspring more than under diploidy [50••]. This suggests that paternal care might be overrepresented in haplodiploids, although empirical support is ambiguous (Table S1). Exclusively paternal care is rare among insects (probably for reasons unrelated to ploidy). The only clear examples are found in three (sub)families of diploid Hemipterans and one family of haplodiploid and strong inbreeding thrips [51], which seem to fit the model well.

One important assumption of these models [48–50••] is that parents are unable to preferentially care for the offspring to which sex they are most related. Under outbreeding, haplodiploid females are equally related to both offspring sexes, but fathers are related only to daughters. Therefore, if males are able to preferentially care for their daughters, paternal care might be promoted. The same might be true for maternal care under inbreeding, as mothers become more related to their daughters than to their sons.

Conclusion

Haplodiploid reproduction is widespread among arthropods. Males either do not inherit any genes from their fathers or, if they do, they fail to pass them on to their offspring. Here we discussed how haplodiploidy can profoundly alter mating system evolution, sex allocation and the evolution of traits under sexual selection or sexual antagonism. These predictions are not just significant to understanding the evolution of haplodiploid taxa, but, in comparison, with diplodiploid taxa, could provide more general insights into these phenomena. Unfortunately, predictions are overwhelmingly based on verbal or very simplistic models, and much of the formal theory that is available was developed for other purposes: either extrapolated from models of X-linkage or motivated by a presumed link with eusociality and therefore tailored specifically to hymenopterans. It is clear that more formal theoretical effort is needed. A particular challenge will be to address how the different types of haplodiploidy (arrhenotoky and PGE with haploid or somatically diploid males) could help dissect the relative importance of the ability of virgin birth and effects of haploid gene expression or transmission. In Table 1, we present how we expect them to affect each of the traits discussed in

the manuscript. As each of the three types of haplodiploidy has evolved repeatedly, these predictions lend themselves well to a formal phylogenetic comparative approach. Data presented in Table S1 could serve as an excellent starting point. Finally, there is scope for a multitude of empirical tests to test predictions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cois.2015.04.018.

Acknowledgements

We would like to thank Deborah Charlesworth, Koos Boomsma, David Shuker, Andy Gardner, Ben Normark and two anonymous reviewers for their valuable comments and encouragement. SAB is supported by a BBSRC Eastbio DTP, AGF by the Darwin Trust and LR by a NERC IRF.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Shuker D, Simmons L: *The Evolution of Insect Mating Systems*. Oxford University Press; 2014.

2. Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman T-L, Hahn MW, Kitano J, Mayrose I, Ming R *et al.*: **Sex determination: why so many ways of doing it?** *PLoS Biol* 2014, **12**:e1001899 Extensive review on the diversity and evolution sex determining systems across different organisms.

3. Normark BB: **Modes of reproduction**. *The Evolution of Insect Mating Systems*. Oxford University Press; 2014:: 1-19.

4. Normark BB: **The evolution of alternative genetic systems in insects**. *Annu Rev Entomol* 2003, **48**:397-423 The most comprehensive review of the genetic/reproductive systems found in insects.

5. Ross L, Gardner A, Hardy N, West SA: **Ecology, not the genetics of sex determination, determines who helps in eusocial populations**. *Curr Biol* 2013, **23**:2383-2387.

6. Gardner A, Ross L: **Mating ecology explains patterns of genome elimination**. *Ecol Lett* 2014, **17**:1602-1612 Theoretical model of the evolution of paternal genome elimination under inbreeding.

7. Werren JH: **The evolution of inbreeding in haplodiploid organisms**. *The Natural History of Inbreeding and Outbreeding: Theoretical and Empirical Perspectives*. University of Chicago Press; 1993.

8. Price GR: **Selection and covariance**. *Nature* 1970, **227**:520-521.

9. Fisher RA: *The Genetical Theory of Natural Selection*. Clarendon Press; 1930.

10. Zahavi A: **Mate selection — a selection for a handicap**. *J Theor Biol* 1975, **53**:205-214.

11. •• Boulton RA, Collins LA, Shuker DM: **Beyond sex allocation: the role of mating systems in sexual selection in parasitoid wasps**. *Biol Rev Cambridge Phil Soc* 2014 http://dx.doi.org/10.1111/brv.12126.
A comprehensive review of research on sexual selection in parasitoid wasps.

12. Reeve HK, Pfennig DW: **Genetic biases for showy males: are some genetic systems especially conducive to sexual selection?** *Proc Natl Acad Sci* 2003, **100**:1089-1094.

13. Kirkpatrick M, Hall DW: **Sexual selection and sex linkage**. *Evolution* 2004, **58**:683-691.

14. Albert AYK: **Sexual selection can resolve sex-linked sexual antagonism**. *Science* 2005, **310**:119-121.

15. •• Andrés JA, Morrow EH: **The origin of interlocus sexual conflict: is sex-linkage important?** *J Evol Biol* 2003, **16**:219-223.
A theoretical study on sex-linked interlocus sexual conflict that is applicable to haplodiploidy.

16. Kraaijeveld K: **Male genes with nowhere to hide; sexual conflict in haplodiploids**. *Anim Biol* 2009, **59**:403-415 Comprehensive review and theory paper on sexual conflict under haplodiploidy.

17. Parker GA, Begon ME: **Sperm competition games: sperm size and number under gametic control**. *Proc R Soc Lond B Biol* 1993, **253**:255-262.

18. •• Immler S: **Sperm competition and sperm cooperation: the potential role of diploid and haploid expression**. *Reproduction* 2008, **135**:275-283.
Review paper that discusses the selective pressures that determine sperm cooperation and conflict under diplodiploidy as well as haplodiploidy.

19. Pizzari T, Foster KR: **Sperm sociality: cooperation, altruism, and spite**. *PLoS Biol* 2008, **6**:e130.

20. Jamieson BGM: *The Ultrastructure and Phylogeny of Insect Spermatozoa*. CUP Archive; 1987.

21. Keller L, Peer K, Bernasconi C, Taborsky M, Shuker DM: **Inbreeding and selection on sex ratio in the bark beetle *Xylosandrus germanus***. *BMC Evol Biol* 2011, **11**:359.

22. Gottlieb D, Lubin Y, Harari AR: **The effect of female mating status on male offspring traits**. *Behav Ecol Sociobiol* 2014, **68**:701-710.

23. Borsa P, Kjellberg F: **Experimental evidence for pseudo-arrhenotoky in *Hypothenemus hampei* (Coleoptera: Scolytidae)**. *Heredity* 1996, **76**:130-135.

24. Werren JH: **The evolution of inbreeding in haplodiploid organisms**. In *The Natural History of Inbreeding and Outbreeding: Theoretical and Empirical Perspectives.* Edited by Thornhill NW. University of Chicago Press; 1993.

25. •• Henter H: **Inbreeding depression and haplodiploidy: experimental measures in a parasitoid and comparisons across diploid and haplodiploid insect taxa**. *Evolution* 2003, **57**:1793-1803.
A comprehensive experimental and comparative analysis of inbreeding depression across insect taxa.

26. Antolin MF: **A genetic perspective on mating systems and sex ratios of parasitoid wasps**. *Res Popul Ecol* 1999, **41**:29-37.

27. Tien NSH, Sabelis MW, Egas M: **Inbreeding depression and purging in a haplodiploid: gender-related effects**. *Heredity* 2015, **114**:327-332.

28. Heimpel GE, de Boer JG: **Sex determination in the hymenoptera**. *Annu Rev Entomol* 2008, **53**:209-230.

29. Sánchez L: **Sciara as an experimental model for studies on the evolutionary relationships between the zygotic, maternal and environmental primary signals for sexual development**. *J Genet* 2010, **89**:325-331.

30. Tabadkani SM, Khansefid M, Ashouri A: **Monogeny, a neglected mechanism of inbreeding avoidance in small populations of gall midges**. *Entomol Exp Appl* 2011, **140**:77-84.

31. Macke E, Magalhães S, Khanh HD-T, Frantz A, Facon B, Olivieri I: **Mating modifies female life history in a haplodiploid spider mite**. *Am Nat* 2012, **179**:E147-E162 http://dx.doi.org/10.1086/665002.

32. Adamson M, Ludwig D: **Oedipal mating as a factor in sex allocation in haplodiploids**. *Phil Trans R Soc Lond B: Biol Sci* 1993, **341**:195-202.

33. Hamilton WD: **Extraordinary sex ratios**. *Science* 1967, **156**:477-488.

34. Jordal BH, Beaver RA, Kirkendall LR: **Breaking taboos in the tropics: incest promotes colonization by wood-boring beetles**. *Global Ecol Biogeogr* 2001, **10**:345-357.

35. Ridley M: **Clutch size and mating frequency in parasitic Hymenoptera**. *Am Nat* 1993, **142**:893-910.

36. Rautiala P, Helanterä H, Puurtinen M: **Unmatedness promotes the evolution of helping more in diplodiploids than in haplodiploids**. *Am Nat* 2014, **184**:318-325.

37. Baer B, Schmid-Hempel P: **Experimental variation in polyandry affects parasite loads and fitness in a bumble-bee**. *Nature* 1999 http://dx.doi.org/10.1038/16451.

38. Oldroyd BP, Fewell JH: **Genetic diversity promotes homeostasis in insect colonies**. *Trends Ecol Evol* 2007, **22**:408-413.

39. Bull JJ: *The Evolution of Sex Determining Mechanisms*. Benjamin Cummings; 1983.

40. Gardner A, Ross L: **Haplodiploidy, sex-ratio adjustment, and eusociality**. *Am Nat* 2013, **181**:E60-E67.

41. Shuker D, Moynihan A, Ross L: **Sexual conflict, sex allocation and the genetic system**. *Biol Lett* 2009, **5**:682.

42. Shuker DM, Cook N: **Evolution: conflict by the sexes, for the sexes**. *Curr Biol* 2014, **24**:R1135-R1137.

43. Gardner A: **Total reproductive value of juvenile females is twice that of juvenile males under X-linkage and haplodiploidy**. *J Theor Biol* 2014, **359**:236-237.

44. Shuker D, Sykes EM, Browning LE, Beukeboom L, West S: **Male influence on sex allocation in the parasitoid wasp *Nasonia vitripennis***. *Behav Ecol Sociobiol* 2006, **59**:829-835.

45. King BH, Kuban KA: **Should he stay or should he go: male influence on offspring sex ratio via postcopulatory attendance**. *Behav Ecol Sociobiol* 2012, **66**:1165-1173.

46. •• Macke E, Olivieri I, Magalhães S: **Local mate competition mediates sexual conflict over sex ratio in a haplodiploid spider mite**. *Curr Biol* 2014, **24**:2850-2854.
Experimental evolution study that provides support that haplodiploid males can alter the sex ratio his mate produces.

47. Ratnieks FLW, Boomsma JJ: **Facultative sex allocation by workers and the evolution of polyandry by queens in social Hymenoptera on JSTOR**. *Am Nat* 1995, **145**:969-993.

48. Wade MJ: **Maternal effect genes and the evolution of sociality in haplo-diploid organisms**. *Evolution* 2001, **55**:453-458.

49. Gardner A: **Evolution of maternal care in diploid and haplodiploid populations**. *J Evol Biol* 2012, **25**:1479-1486
Theoretical model exploring how haplodiploidy can affect the evolution of maternal care.

50. •• Davies NG, Gardner A: **Evolution of paternal care in diploid and haplodiploid populations**. *J Evol Biol* 2014, **27**:1012-1019.
Theoretical model exploring how haplodiploidy can affect the evolution of paternal care.

51. Tallamy DW: **Evolution of exclusive paternal care in arthopods**. *Annu Rev Entomol* 2001, **46**:139-165.

52. Normark BB: **Unusual gametic and genetic systems**. In *Sperm Biology: An Evolutionary Perspective.* Edited by Hosken DJ, Birkhead T. Academic Press; 2009.

53. Wells HG, Huxley JS, Wells GP: *The Science of Life*. Amalgamated Press; 1929.

54. Varndell N, Godfray H: **Facultative adjustment of the sex ratio in an insect (*Planococcus citri*, Pseudococcidae) with paternal genome loss**. *Evolution* 1996, **50**:2100-2105.

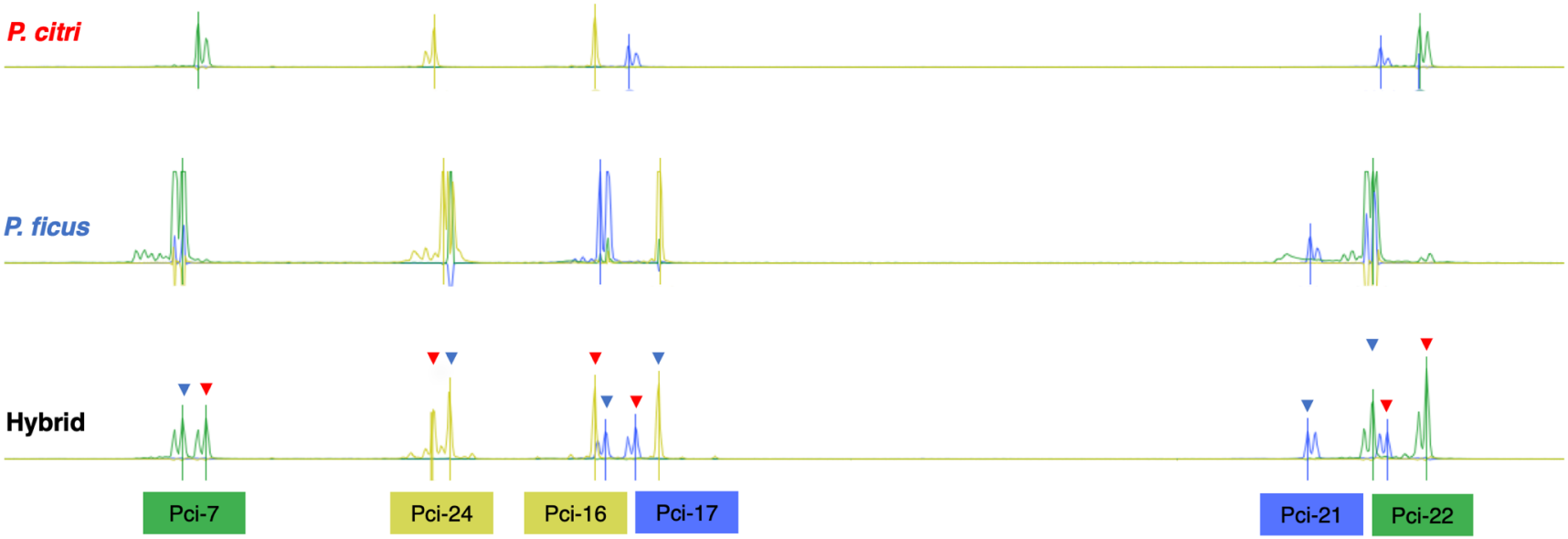
55. Ross L, Langenhof M, Pen I, Beukeboom L, West S, Shuker D: **Sex allocation in a species with paternal genome elimination: the roles of crowding and female age in the mealybug *Planococcus citri***. *Evol Ecol Res* 2010, **12**:89-104.

56. •• Featherston R, Jones TM, Elgar MA: **Female resistance behaviour and progeny sex ratio in two Bradysia species (Diptera: Sciaridae) with paternal genome elimination**. *J Evol Biol* 2013, **26**:919-928.
One of only a few experimental studies on the reproductive ecology of a species with PGE.

Appendix 2

Supplementary information for Chapter 2

Supplementary Figure S1. An example of microsatellite loci and species-diagnostic alleles in a hybrid family. Top, *P. citri* mother. Middle, *P. ficus* father. Bottom, F1 hybrid (maternally-inherited alleles are highlighted with red arrows, paternally-inherited alleles with blue arrows).



Supplementary Table S1. Microsatellite loci (Martins *et al.* 2012) used in this study: allele richness (N), allele size range and genomic location in our *P. citri* and *P. ficus* experimental populations.

Locus	***Planococcus citri***			***Planococcus ficus***		
	N	*Alleles*	*Location**	*N*	*Alleles*	*Location**
Pci-7	3	137, 140, 143	00125 (F 0.01 / R 0.004)	4	137, 140, 143, 146	09272 (F 0.01 / R 0.004)
Pci-16	2	191, 194	01444 (F 0.004 / R 0.01)	1	200	09756 (F 0.004 / R 0.01)
Pci-17	2	197, 200	00790 (F 0.0004 / R 0.02)	1	194	08745 (F 0.02 / R 0.02)
Pci-21	1	288	00133 (F 0.01 / R 0.004)	1	279	01154 (F 0.01 / R 0.004)
Pci-22	1	295	00083 (F 0.01 / R 0.0001)	2	289, 292	00369 (F 0.6 / R 0.0001)
Pci-24	1	168	00250 (F 0.00 / R 0.004)	1	172	03585 (F 0.001 / R 0.004)

* For genomic locations of each locus, scaffold numbers corresponding to best BLAST hits of primer sequences to assemblies PCITRI.V1 and PFICUS.V0 are given. All forward (F) and reverse (R) pairs had best hits to the same scaffold in all loci; E-values are indicated in superscript.

Supplementary Table S2. Paternal allele escapes found in the present study. Paternal alleles in F1 and escaped genotypes in F2 are shown in bold. Paternal TR, paternal transmission ratio

F1 male	Locus	F1 male genotype	F1 female genotype	F2	F2 genotypes	Paternal TR
W1-4_1	*Pci-17*	**194**/197	197/200	12	5 197/197, 6 197/200, 1 **194/197**	0.917
	Pci-7	**140**/143	140/140	11	11 140/143	1
	Pci-16	191/**200**	191/191	12	12 191/191	1
	Pci-21	**279**/288	288/288	12	12 288/288	1
	Pci-22	**292**/295	295/295	12	11 295/295, **1 292/295**	0.917
	Pci-24	168/**172**	168/168	12	12 168/168	1
BC-1_3	*Pci-17*	**197**/200	197/200	11	7 197/200, 4 200/200	1
	Pci-7	140/**143**	?	11	10 140/140, 1 **140/143**	0.909*
	Pci-16	**191**/194	194/194	11	11 194/194	1
BW-2_3	*Pci-17*	197/197	197/197	12	12 197/197	–
	Pci-7	**137**/143	143/143	12	11 143/143, 1 **137/143**	0.917
	Pci-16	191/191	191/191	12	12 191/191	–
CB-3_1	*Pci-17*	197/**200**	197/197	12	11 197/197, 1 197/200	0.917
	Pci-7	**140**/143	143/143	12	11 143, 143, 1 140/143	0.917
	Pci-16	191/**194**	194/194	12	12 191/194	1

Supplementary Table S3. Summary of fitted models

	Model 1 **Sex pheromone response (all males)**			**Model 2** **Total time in contact (responding males)**			**Model 3** **Time to first contact (responding males)**		
Response Variable	I(contact.time != 0)			logit(total contact time/900)			logit(time to first contact/900)		
Fixed Parts	*Odds Ratio*	*CI*	*p*	*B*	*CI*	*p*	*B*	*CI*	*p*
(Intercept)	8.09	2.94 – 22.29	<.001	0.02	-0.61 – 0.64	0.959	-1.21	-1.78 – -0.64	<.001
pheromoneficus	0.11	0.04 – 0.33	<.001	-2.28	-3.19 – -1.36	<.001	0.44	-0.35 – 1.23	0.277
genotypef2	0.81	0.26 – 2.48	0.709	0.08	-0.67 – 0.83	0.834	-0.47	-1.17 – 0.23	0.187
genotypeficus	0.15	0.04 – 0.56	0.004	-1.34	-2.50 – -0.19	0.023	0.42	-0.67 – 1.50	0.454
orderFC	1	0.52 – 1.90	0.997	0.12	-0.44 – 0.69	0.668	0.26	-0.28 – 0.80	0.352
pheromoneficus:genotypef2	2.18	0.57 – 8.42	0.257	0.32	-0.87 – 1.50	0.601	0.5	-0.57 – 1.58	0.358
pheromoneficus:genotype ficus	41.53	6.51 – 265.15	<.001	2.76	1.23 – 4.30	<.001	-0.81	-2.17 – 0.56	0.248
Random Parts									
N_{ID}		143			127			119	
ICC_{ID}		0.186			0.109			0.195	
Observations		288			199			178	

Supplementary Table S4. Species confirmation

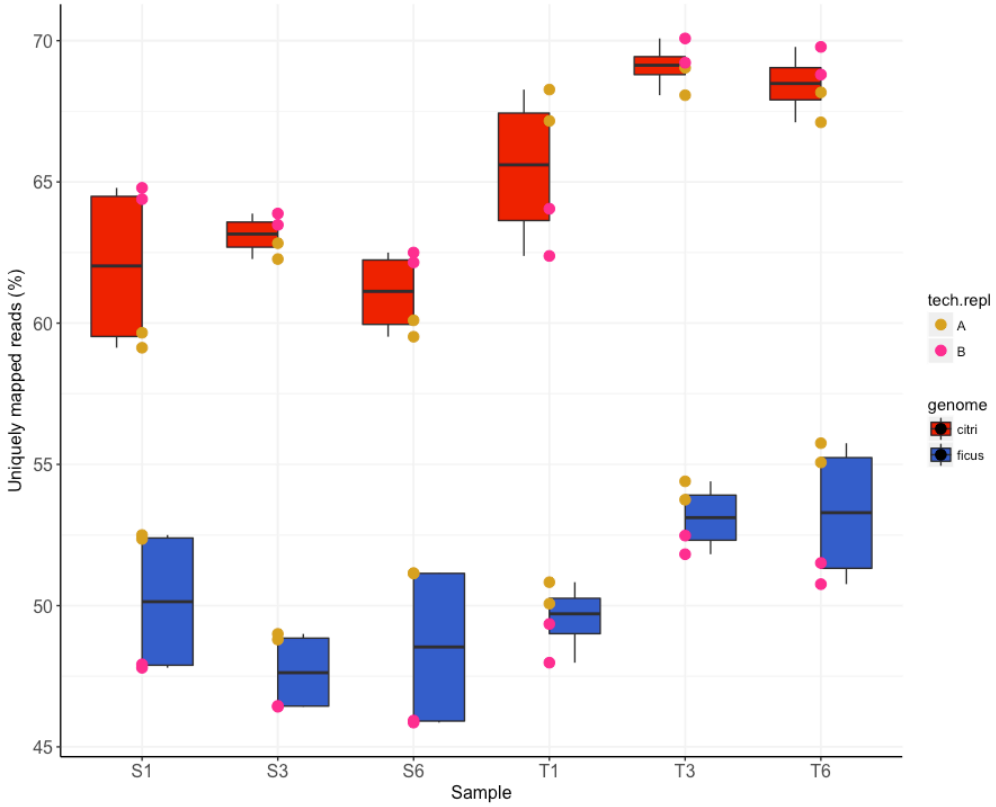
Region	**Accession***	**Assembly**	**Best hit to assembly**	**Best GenBank hit**		
				Accession	*Cover / Identity / E-value*	*Species*
28S–D2	GU134660	PCITRI.V1	No hit	-	-	-
		PFICUS.V0	PFICUS_24458:1772-2081	AY427341.1	100% / 100% / 1e-159	*P. ficus*
ITS2	GU134678	PCITRI.V1	PCITRI_07258:6707-7394	HE984349.1	100% / 99% / 0.0	*P. citri*
		PFICUS.V0	PFICUS_24458:648-1339	GU134677.1	100% / 100% / 0.0	*P. ficus*
COI–region 2	GU134691	PCITRI.V1	PCITRI_05460:11187-11366	MF952478.1	100% / 99% / 2e-84	*P. citri*
		PFICUS.V0	PFICUS_56406:570-789	GU134677.1	100% / 99% / 8e-105	*P. ficus*
COI–LCO	GU134705	PCITRI.V1	PCITRI_00260:3357-3170	MG813768.1	100% / 91% / 7e-65	*P. citri*
		PFICUS.V0	PFICUS_56406:60-535	LC190452.1	100% / 99% / 0.0	*P. ficus*

* From Malausa *et al.* (2011)

Appendix 3

Supplementary information for Chapter 3

Supplementary Figure S1. Percentage of uniquely mapped RNA-seq reads to PCITRI.V0 (red) and PFICUS.V0 (blue). Each dot corresponds to reads generated in a single lane for each sequencing library. S = soma, T = testis. Orange = technical replicate A, pink = technical replicate B.



Supplementary Table S1. PCITRI.V0 and PFICUS.V0 assembly and annotation stats

Metric	**PCITRI.v0**	**PFICUS.v0**
Span (bp)	486,392,826	334,968,598
N50 (bp)	124,550	9,321
GC (%)	34.4	34.3
Longest scaffold (bp)	1,541,683	84,905
Scaffolds (N)	17,212	66,754
Ns (%)	4.93	0.02
Repeats (%)	2.61%	2.86%
Genes (N)	39,801	33,169
Transcripts (N)	41,192	26,694
BUSCO complete (%)	96.7	94.4
BUSCO duplicated (%)	6.9	4.6
BUSCO fragmented (%)	1.0	2.6

Supplementary Table S2. Raw and filtered ASE sites found in F1 soma and testis RNA-seq

ASE sites	**Between replicates**		**Shared by all replicates**		**Final set (after filters)**	
Tissue	**Soma**	**Testis**	**Soma**	**Testis**	**Soma**	**Testis**
Total	269,232	395,777	179,212	251,730	170,916	243,193
Exonic	133,150	145,739	102,045	111,726	94,758	104,774
Intronic	43,270	84,089	21,790	42,784	21,555	42,364
Intergenic	90,233	158,876	54,003	93,701	53,240	92,557
Orphan	2,579	7,073	1,374	3,519	1,363	3,498

Supplementary Table S3. Summary statistics of read depth and expression bias to maternal genome. Mean, standard deviation and median for each annotation feature are indicated.

ASE sites	**Read depth**						**Expression bias to maternal genome**					
Tissue	**Soma**			**Testis**			**Soma**			**Testis**		
	Mean	*SE*	*Med*	*Mean*	*SD*	*Med*	*Mean*	*SD*	*Med*	*Mean*	*SD*	*Med*
Exonic	231.4	605.5	94.3	323.8	834.5	115	0.873	0.120	0.900	0.96	0.06	0.976
Intronic	81.2	167.4	64.3	115.5	286.9	67	0.921	0.116	0.963	0.99	0.05	0.997
Intergenic	168.7	605.6	55.3	192.1	639.2	75	0.927	0.101	0.967	0.98	0.06	0.993
Orphan	139.8	464.7	55.3	137.2	297.3	66	0.941	0.098	0.973	0.98	0.08	0.997

Supplementary Table S4. Paternal only (P), paternally-biased (PB) and biparental (B) genes in soma. Bias to the maternal genome (p_m), number of informative SNPs (#), bias in testis and contig they are situated on (loc) are indicated. Reciprocal orthologues in *D. melanogaster* (and tissue-specific enrichment, if available) and *A. pisum* and gene function, as inferred by BLASTp, GO terms and InterPro domains, are shown.

Gene	Bias	p_m	#	Bias in T	Loc	In *D. mel*	In *A. pisum*	Function
g38721	P	0	1	P	00335	No hit	No hit	Unknown
g20597	P	0.04	3	-	02436	No hit	No hit	Unknown
g29917	PB	0.09	5	-	06730	No hit	No hit	Unknown
g6888	PB	0.09	42	P	00792	No hit	No hit	Peptidase aspartic, putative (retrotransposon?)
g32912	PB	0.1	5	-	00093	No hit	No hit	Unknown
g17372	PB	0.14	11	-	01839	No hit	No hit	Beta-ketoacyl synthase (fatty acid synthesis)
g8434	PB	0.14	9	-	00915	No hit	No hit	Retrotransposon?
g17368	PB	0.16	6	-	01839	No hit	No hit	Beta-ketoacyl synthase (fatty acid synthesis)
g19257	PB	0.16	2	-	02173	No hit	No hit	Glyceraldehyde 3-phosphate dehydrogenase
g23018	PB	0.17	3	-	03083	No hit	No hit	Unknown
g8627	PB	0.2	10	-	00936	No hit	No hit	Polyketide synthase (fatty acid synthesis)
g8626	PB	0.21	2	-	00936	No hit	No hit	Beta-ketoacyl synthase (fatty acid synthesis)
g36032	PB	0.22	13	-	00207	No hit	No hit	Polyketide synthase (fatty acid synthesis)
g1013	PB	0.23	8	-	00018	No hit	No hit	Carbonic anhydrase
g2208	PB	0.24	6	-	00456	No hit	ACYPI004105	Vitamin K epoxide reductase
g7112	PB	0.24	11	PB	00809	No hit	No hit	Beta-ketoacyl synthase (fatty acid synthesis)
g33031	PB	0.26	9	PB	00098	FBgn0030334 (Karl) *Carcass*	ACYPI008487	Unknown
g36031	PB	0.27	4	-	00207	No hit	No hit	Beta-ketoacyl synthase (fatty acid synthesis)
g10367	PB	0.28	17	M	01094	FBgn0053548 (SmydA-8) *Head, eye*	ACYPI44442	SET domain, histone lysine methylation
g4366	PB	0.28	8	-	00594	No hit	ACYPI46979	Unknown
g14309	PB	0.31	3	B	00035	No hit	No hit	Unknown
g98	PB	0.32	2	-	00001	No hit	No hit	Unknown

g2935	PB	0.33	9	-	00500	No hit	No hit	Pentatricopeptide repeat-containing protein (mitochondrial?)
g29936	PB	0.33	2	-	06761	No hit	No hit	Unknown
g34746	PB	0.33	5	B	00150	No hit	No hit	Unknown
g36770	PB	0.33	16	MB	00239	FBgn0032053 (mRpL51) *Testis*	ACYPI006343	39S ribosomal protein L51 (mitochondrial)
g8624	PB	0.33	21	-	00936	No hit	No hit	Beta-ketoacyl synthase (fatty acid synthesis)
g22669	PB	0.35	20	-	02973	No hit	No hit	ABC transporter (ATPase activity)
g7111	PB	0.35	33	-	00809	No hit	No hit	Beta-ketoacyl synthase (fatty acid synthesis)
g37995	PB	0.36	8	MB	00298	FBgn0030460 (CG2453)	ACYPI005028	Ubiquinone/menaquinone biosynthesis methyltransferase ubiE (mitochondrial)
g11998	PB	0.37	3	-	01276	No hit	No hit	Unknown
g21159	PB	0.38	4	-	02570	No hit	No hit	Unknown
g222	PB	0.38	12	MB	00003	FBgn0031660 (mRpL28)	ACYPI071124	39S ribosomal protein L28 (mitochondrial)
g32866	PB	0.38	43	MB	00092	FBgn0052649 (CG32649)	ACYPI004004	UbiB domain, protein kinase (mitochondrial?)
g23738	PB	0.39	10	MB	03323	No hit	No hit	Long-chain acyl-CoA dehydrogenase (mitochondrial)
g10234	PB	0.4	19	-	01084	FBgn0025885 (Inos)	ACYPI006105	Myo-inositol-1-phosphate synthase (glucose metabolism)
g2099	PB	0.4	47	MB	00450	No hit	ACYPI001597	Heat shock protein
g28199	PB	0.4	43	-	00056	FBgn0033816 (CG4679) *Hindgut*	ACYPI001244	Pentatricopeptide repeat-containing ribosomal protein (mitochondrial)
g37590	PB	0.41	11	-	00277	FBgn0034727 (mRpS29)	ACYPI002770	28S ribosomal protein S29 (mitochondrial)
g38698	PB	0.41	17	MB	00334	FBgn0003071 (Pfk) Testis	ACYPI001940	ATP-dependent phosphofructokinase activity
g4839	PB	0.47	5	MB	00634	No hit	No hit	ATP synthesis (mitochondrial)
g15069	PB	0.48	2	MB	01512	No hit	No hit	Phosphate substrate carrier (mitochondrial)
g15314	B	0.41	2	-	01545	No hit	No hit	ABC transporter (ATPase activity)
g33524	B	0.42	4	-	00110	No hit	No hit	MFS transporter
g35997	B	0.43	6	-	00206	No hit	No hit	MFS transporter
g25371	B	0.44	3	-	04006	FBgn0030786 (mRpL22) *Hindgut*	ACYPI002040	39S ribosomal protein L22 (mitochondrial)
g28606	B	0.46	2	-	00067	FBgn0033883 (CG16935)	ACYPI001135	Trans-2-enoyl-CoA reductase (mitochondrial fatty acid synthesis?)
g38200	B	0.47	7	-	00308	FBgn0037608 (mRpL19)	ACYPI004065	39S ribosomal protein L19 (mitochondrial)

g5334	B	0.47	2	-	00668	FBgn0032031 (CG13390)	ACYPI006444	GTPase (mitochondrial)
g12633	B	0.48	12	-	01352	FBgn0001995 (mRpL4) *TA ganglion, hindgut*	ACYPI008888	39S ribosomal protein L4 (mitochondrial)
g22177	B	0.48	2	MB	02827	No hit	No hit	Carbohydrate kinase
g31695	B	0.49	2	B	10660	No hit	No hit	Long chain fatty acid Co-A ligase
g35667	B	0.49	3	-	00189	No hit	ACYPI32743	39S ribosomal protein L10 (mitochondrial)
g4720	B	0.49	8	-	00624	FBgn0033100 (CG3420)	ACYPI009248	CDGSH iron sulphur-containing domain (mitochondrial)
g5142	B	0.49	10	-	00654	No hit	No hit	Ligand-gated ion channels
g16558	B	0.5	6	-	01717	FBgn0033215 (CG1942) *Midgut*	ACYPI000410	Acyltransferase
g7114	B	0.5	2	-	00809	No hit	No hit	Polyketide synthase (fatty acid synthesis)
g35412	B	0.51	4	-	00179	FBgn0002526 (LanA)	ACYPI010019	Laminin subunit
g38423	B	0.51	29	MB	00318	FBgn0030692 (mRpS30) *Hindgut*	ACYPI007089	28S ribosomal protein S30 (mitochondrial)
g6978	B	0.51	20	M	00799	FBgn0039689 (CIA30)	ACYPI005395	NADH:ubiquinone oxidoreductase (mitochondrial)
g8411	B	0.51	11	-	00914	FBgn0030552 (mRpL38) *TA ganglion*	ACYPI008716	39S ribosomal protein L38 (mitochondrial)
g9728	B	0.51	10	-	01036	No hit	No hit	Unknown
g14179	B	0.52	12	MB	00032	FBgn0033879 (CG6543) *M. tubules*	ACYPI001292	Enoyl-CoA hydratase/isomerase (mitochondrial)
g16297	B	0.52	1	-	01681	No hit	No hit	Unknown
g24280	B	0.52	3	-	03519	No hit	No hit	Ligand-gated ion channels
g3639	B	0.52	3	-	00545	FBgn0031500 (CG17221) *Acc. glands, carcass*	ACYPI009182	Alcohol dehydrogenase (mitochondrial)
g13384	B	0.53	12	-	01439	FBgn0036857 (CG16935)	ACYPI000430	Enoyl-[acyl-carrier-protein] reductase (mitochondrial fatty acid synthesis)
g1855	B	0.53	2	-	00438	No hit	No hit	FAST kinase (mitochondrial)
g28605	B	0.53	2	-	00067	No hit	No hit	Unknown
g31499	B	0.53	2	-	10020	No hit	No hit	WWE domain (protein-protein interactions)
g14776	B	0.54	2	-	00047	No hit	No hit	Unknown
g16351	B	0.54	2	-	01690	No hit	No hit	Unknown
g17010	B	0.54	12	-	01784	FBgn0261800 (LanB1) *Crop*	ACYPI088486	Laminin

g12149	B	0.55	4	-	01293	No hit	No hit	Short-chain dehydrogenase/reductase
g914	B	0.55	11	-	00016	FBgn0039184 (CG6432) *Carcass*	ACYPI000901	Acyl-CoA synthetase (fatty acid synthesis (mitochondrial)?)
g2251	B	0.56	4	-	00460	No hit	No hit	Unknown
g33957	B	0.56	3	M	00123	FBgn0035772 (Sh3) *Midgut, acc. glands*	ACYPI008600	SH3-binding, glutamic acid-rich protein
g8251	B	0.56	4	M	00898	FBgn0038876 (CG5919)	ACYPI007594	Isopentenyl diphosphate isomerase (mitochondrial)
g35763	B	0.57	1	MB	00193	FBgn0052264 (CG32264) *TA ganglion*	ACYPI007330	RPEL repeat (actin binding)
g26248	B	0.58	2	-	04495	No hit	No hit	Glycosyl transferase (protein kinase)
g5737	B	0.58	4	-	00704	No hit	No hit	Unknown
g11770	B	0.6	1	M	01249	No hit	No hit	SH3 domain, immunoglobulin
g3175	B	0.6	2	-	00516	FBgn0032644 (CG5131)	ACYPI008652	ATP23, peptidase (mitochondrial)
g5062	B	0.61	3	-	00648	No hit	No hit	UDP glycosyltransferase
g14378	B	0.62	2	-	00036	FBgn0050499 (CG30499) *Testis*	ACYPI003993	Ribulose-phosphate 3-epimerase
g8055	B	0.62	2	MB	00885	No hit	ACYPI007177	Short-chain dehydrogenases/reductase
g15028	B	0.63	2	-	01508	No hit	No hit	MFS transporter
g37880	B	0.63	2	MB	00294	FBgn0037809 (CG12818) *Testis*	ACYPI006162	CBF1-interacting co-repressor CIR
g14705	B	0.64	2	-	00044	No hit	No hit	BTB/POZ domain
g32183	B	0.64	2	M	13236	No hit	No hit	Unknown
g38000	B	0.65	2	MB	00298	FBgn0033367 (PPO2)	ACYPI072244	Prophenoloxidase (extracellular matrix?)

Supplementary Table S5. Paternal only (P), paternally-biased (PB) and biparental (B) genes in testis. Bias to the maternal genome (p_m), number of informative SNPs (#), bias in soma and contig they are situated on (loc) are indicated. Reciprocal orthologues in *D. melanogaster* and *A. pisum* and gene function, as inferred by BLASTp, GO terms and InterPro domains, are shown.

Gene	Bias	p_m	#	Bias in S	Loc	In *D. mel*	In *A. pisum*	Function
g38721	P	0	1	P	00335	No hit	No hit	Unknown
g6888	P	0.02	6	PB	00792	No hit	No hit	Peptidase aspartic, putative (retrotransposon?)
g21988	PB	0.15	12	-	02781	No hit	No hit	Unknown
g23345	PB	0.16	1	-	03209	FBgn0031068 (Air)	ACYPI28318	ERV/ALR sulfhydryl oxidase domain (mitochondrial)
g26531	PB	0.23	18	-	04685	No hit	No hit	Membrane metallo-endopeptidase-like 1
g10734	PB	0.25	11	-	01134	No hit	No hit	Unknown
g7112	PB	0.26	2	PB	00809	No hit	No hit	Beta-ketoacyl synthase (fatty acid synthesis)
g33031	PB	0.27	7	PB	00098	FBgn0030334 (CG4139)	ACYPI008487	Unknown
g35539	PB	0.34	5	-	00184	No hit	No hit	Alpha crystallin/Hsp20 domain (heat shock protein)
g22380	PB	0.38	3	-	02892	No hit	No hit	Laminin
g37928	PB	0.38	4	-	00295	No hit	No hit	Unknown
g19277	PB	0.45	6	-	02175	No hit	No hit	Serine proteases, trypsin domain (proteolysis)
g14309	B	0.42	3	PB	00035	No hit	No hit	Unknown
g31695	B	0.46	2	B	10660	No hit	No hit	Long chain fatty acid Co-A ligase
g34746	B	0.57	3	PB	00150	No hit	No hit	Unknown
g39590	B	0.57	6	-	00379	No hit	No hit	Facilitated trehalose transporter
g37105	B	0.59	3	MB	00254	FBgn0003137 (Papilin)	ACYPI003244	Thrombospondin
g666	B	0.59	3	-	00011	No hit	ACYPI072010	Zinc finger protein

Appendix 4

Supplementary information for Chapter 4

Supplementary Table S1. Paternal and maternal transmission ratios for all families and loci.

#F1, number of unambiguous F1 genotypes for each locus.

Pat. A/Pat. B, counts of most common and alternative paternal alleles in F1 genotypes.

Mat. A/Mat. B, counts of both maternal alleles in F1 genotypes chosen randomly.

Pat. TR/Mat. TR, transmission ratio of Pat. A/Mat. A.

P, exact binomial test probability associated between observed and expected counts of Pat. A/Mat. A in F1 under Mendelian expectations. Significant deviations at the $\alpha = 0.01$ level are highlighted in bold and indicated with two asterisks. Not significant deviations at 0.01 level that are significant at the conventional 0.05 level are indicated with a single asterisk. NS, not significant at any of these significance level.

Family	Locus	# F1	Pat. A	Pat. B	Pat. TR	P		Mat. A	Mat. B	Mat. TR	P	
Head louse												
LFH_02	M3_10	7	7	0	1.000	0.016	*	-	-	-	-	
	M2_19	8	8	0	1.000	**0.008**	**	-	-	-	-	
	T2_6	8	-	-	-	-		5	3	0.375	0.727	NS
	T4_5	9	9	0	1.000	**0.004**	**	-	-	-	-	
LFH_03	M3_10	10	10	0	1.000	**0.002**	**	-	-	-	-	
	M3_19	10	10	0	1.000	**0.002**	**	-	-	-	-	
	M2_2	8	8	0	1.000	**0.008**	**	5	3	0.375	0.727	NS
	M2_19	8	-	-	-	-		6	2	0.750	0.289	NS
	T2_7	8	-	-	-	-		5	3	0.375	0.727	NS
	T4_5	8	8	0	1.000	**0.008**	**	4	4	0.500	1.000	NS
LFH_04	M3_10	6	-	-	-	-		3	3	0.500	1.000	NS
	M3_19	7	7	0	1.000	0.016	*	-	-	-	-	
	T2_6	8	8	0	1.000	**0.008**	**	-	-	-	-	
	T1_4	8	-	-	-	-		5	3	0.625	0.727	NS
LFH_05	M3_19	8	-	-	-	-		5	3	0.375	0.727	NS
	T4_5	10	-	-	-	-		6	4	0.600	0.754	NS
	T1_4	10	10	0	1.000	**0.002**	**	-	-	-	-	
LFH_06	M3_19	9	9	0	1.000	**0.004**	**	6	3	0.667	0.508	NS
	M2_3	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	5	5	0.500	1.000	NS
	T1_4	10	-	-	-	-		6	4	0.400	0.754	NS
LFH_07	M3_19	10	-	-	-	-		7	3	0.700	0.344	NS
	M2_2	10	-	-	-	-		5	5	0.500	1.000	NS
	T2_7	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	7	3	0.300	0.344	NS
LFH_08	M3_19	8	8	0	1.000	**0.008**	**	-	-	-	-	
	M2_2	8	8	0	1.000	**0.008**	**	-	-	-	-	
LFH_09	M3_19	9	9	0	1.000	**0.004**	**	-	-	-	-	
	M2_2	9	-	-	-	-		5	4	0.444	1.000	NS
	T2_7	6	6	0	1.000	0.031	*	5	1	0.833	0.219	NS
	T4_5	8	-	-	-	-		4	4	0.500	1.000	NS
	T1_4	10	-	-	-	-		10	0	1.000	**0.002**	**
LFH_10	M3_10	10	-	-	-	-		5	5	0.500	1.000	NS
	M3_19	10	10	0	1.000	**0.002**	**	6	4	0.600	0.754	NS
	M2_2	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	9	9	0	1.000	**0.004**	**	5	4	0.556	1.000	NS
LFH_11	M3_10	10	-	-	-	-		8	2	0.200	0.109	NS
	M3_19	11	11	0	1.000	**0.001**	**	8	3	0.727	0.227	NS
	M2_2	9	9	0	1.000	**0.004**	**	-	-	-	-	
	T4_5	12	12	0	1.000	**< 0.001**	**	6	6	0.500	1.000	NS
	T1_4	11	-	-	-	-		10	1	0.091	0.012	*
LFH_12	M3_10	10	-	-	-	-		7	3	0.700	0.344	NS
	M3_19	9	9	0	1.000	**0.004**	**	6	3	0.333	0.508	NS
	T2_7	10	-	-	-	-		6	4	0.600	0.754	NS
	T4_5	10	10	0	1.000	**0.002**	**	6	4	0.400	0.754	NS
	T1_4	10	10	0	1.000	**0.002**	**	5	5	0.500	1.000	NS
LFH_14	M2_2	8	8	0	1.000	**0.008**	**	6	2	0.250	0.289	NS
	M2_19	7	-	-	-	-		4	3	0.571	1.000	NS
	T4_5	8	-	-	-	-		7	1	0.125	0.070	NS
LFH_15	M3_10	10	10	0	1.000	**0.002**	**	7	3	0.700	0.344	NS
	M3_19	10	-	-	-	-		8	2	0.200	0.109	NS
	M2_2	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	9	1	0.100	0.021	*
	T1_4	10	10	0	1.000	**0.002**	**	-	-	-	-	
LFH_16	M3_10	10	10	0	1.000	**0.002**	**	6	4	0.400	0.754	NS
	T2_6	10	10	0	1.000	**0.002**	**	6	4	0.600	0.754	NS
	T2_7	7	-	-	-	-		4	3	0.429	1.000	NS
	T4_5	9	-	-	-	-		6	3	0.667	0.508	NS
	T1_4	9	9	0	1.000	**0.004**	**	6	3	0.667	0.508	NS
LFH_19	M2_2	10	10	0	1.000	**0.002**	**	7	3	0.700	0.344	NS
	T1_4	10	10	0	1.000	**0.002**	**	6	4	0.400	0.754	NS
LFH_20	M3_19	10	10	0	1.000	**0.002**	**	-	-	-	-	
	M2_2	10	-	-	-	-		7	3	0.300	0.344	NS
	T2_6	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T2_7	8	-	-	-	-		4	4	0.500	1.000	NS
	T1_4	10	9	1	0.900	0.021	*	-	-	-	-	
LFH_21	M3_10	10	-	-	-	-		8	2	0.200	0.109	NS
	M3_19	10	-	-	-	-		7	3	0.700	0.344	NS
	M2_2	9	9	0	1.000	**0.004**	**	5	4	0.444	1.000	NS
	M2_19	8	8	0	1.000	**0.008**	**	-	-	-	-	
	T2_7	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	6	4	0.600	0.754	NS
LFH_24	M2_2	10	-	-	-	-		5	5	0.500	1.000	NS
	M2_3	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T2_7	10	10	0	1.000	**0.002**	**	6	4	0.400	0.754	NS
	T4_5	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T1_4	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T1_4	10	-	-	-	-		6	4	0.600	0.754	NS
LFH_25	M3_19	10	10	0	1.000	**0.002**	**	8	2	0.800	0.109	NS
	T2_6	9	9	0	1.000	**0.004**	**	-	-	-	-	
	T2_7	10	10	0	1.000	**0.002**	**	6	4	0.600	0.754	NS
	T1_4	10	-	-	-	-		6	4	0.400	0.754	NS
LFH_26	M3_19	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	-	-	-	-	
LFH_28	M2_2	10	10	0	1.000	**0.002**	**	5	5	0.500	1.000	NS
	T4_5	10	10	0	1.000	**0.002**	**	6	4	0.400	0.754	NS
	T1_4	10	-	-	-	-		7	3	0.344	0.344	NS
LFH_29	M3_10	10	-	-	-	-		6	4	0.600	0.754	NS
	M3_19	10	-	-	-	-		6	4	0.600	0.754	NS
	M2_2	10	-	-	-	-		7	3	0.300	0.344	NS
	T2_7	10	10	0	1.000	**0.002**	**	5	5	0.500	1.000	NS
	T4_5	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T1_4	10	-	-	-	-		6	4	0.600	0.754	NS
LFH_30	M3_10	10	-	-	-	-		8	2	0.200	0.109	NS
	M3_19	10	10	0	1.000	**0.002**	**	-	-	-	-	
	M2_2	10	-	-	-	-		6	4	0.400	0.754	NS
	T2_6	9	9	0	1.000	**0.004**	**	-	-	-	-	
	T2_7	10	-	-	-	-		7	3	0.300	0.344	NS
LFH_31	M3_10	10	-	-	-	-		7	3	0.700	0.344	NS
	M3_19	10	-	-	-	-		5	5	0.500	1.000	NS
	M2_2	9	9	0	1.000	**0.004**	**	5	4	0.556	1.000	NS
	T2_6	10	10	0	1.000	**0.002**	**	5	5	0.500	1.000	NS
	T2_7	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T1_4	10	-	-	-	-		6	4	0.600	0.754	NS
LFH_33	M3_10	9	-	-	-	-		5	4	0.444	1.000	NS
	M2_2	10	9	1	0.900	0.021	*	6	4	0.600	0.754	NS
	T2_6	10	10	0	1.000	**0.002**	**	7	3	0.300	0.344	NS
	T2_7	9	9	0	1.000	**0.004**	**	5	4	0.556	1.000	NS
	T4_5	9	9	0	1.000	**0.004**	**	7	2	0.333	0.180	NS
	T1_4	9	-	-	-	-		6	3	0.667	0.508	NS
LFH_34	M2_2	7	7	0	1.000	0.016	*	5	2	0.286	0.453	NS
	T2_6	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T2_7	10	9	1	0.900	0.021	*	6	4	0.400	0.754	NS
	T4_5	10	10	0	1.000	**0.002**	**	6	4	0.600	0.754	NS
	T1_4	10	-	-	-	-		7	3	0.300	0.344	NS
Body louse												
LFB_01	M3_10	5	5	0	1.000	0.063	NS	-	-	-	-	
	M3_19	5	-	-	-	-		3	2	0.400	1.000	NS
	M2_13	5	4	1	0.800	0.375	NS	-	-	-	-	
	T4_5	6	6	0	1.000	0.031	*	-	-	-	-	
LFB_02	M2_3	9	-	-	-	-		6	3	0.667	0.508	NS
	T4_5	10	10	0	1.000	**0.002**	**	-	-	-	-	
LFB_04	M3_19	7	-	-	-	-		4	3	0.571	1.000	NS
	M2_19	7	-	-	-	-		6	1	0.143	0.125	NS
	M2_3	7	7	0	1.000	0.016	*	-	-	-	-	
	T4_5	8	-	-	-	-		5	3	0.375	0.727	NS
LFB_06	M3_10	10	-	-	-	-		6	4	0.600	0.754	NS
	M3_19	10	10	0	1.000	**0.002**	**	-	-	-	-	
	M2_19	7	7	0	1.000	0.016	*	-	-	-	-	
	M2_3	10	-	-	-	-		7	3	0.300	0.344	NS
	T2_6	10	7	3	0.700	0.344	NS	-	-	-	-	
	T4_5	18	14	4	0.778	0.031	*	-	-	-	-	
LFB_07	M3_19	8	-	-	-	-		5	3	0.625	0.727	NS
	M2_3	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	6	4	0.600	0.754	NS
LFB_08	M3_10	10	-	-	-	-		8	2	0.200	0.109	NS
	M2_2	9	-	-	-	-		6	3	0.667	0.508	NS
	T2_6	10	-	-	-	-		6	4	0.400	0.754	NS
	T4_5	10	-	-	-	-		7	3	0.700	0.344	NS
LFB_09	M3_10	10	10	0	1.000	**0.002**	**	6	4	0.400	0.754	NS
	T4_5	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T2_6	10	5	5	0.500	1.000	NS	-	-	-	-	
	M2_19	7	-	-	-	-		4	3	0.571	1.000	NS
LFB_10	M3_10	10	10	0	1.000	**0.002**	**	6	4	0.400	0.754	NS
	M3_19	6	6	0	1.000	0.031	*	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	-	-	-	-	
LFB_14	M3_10	10	-	-	-	-		5	5	0.500	1.000	NS
	M3_19	9	-	-	-	-		6	3	0.333	0.508	NS
	M2_2	10	10	0	1.000	**0.002**	**	-	-	-	-	
	M2_3	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	5	5	0.500	1.000	NS
LFB_15	M3_10	9	-	-	-	-		5	4	0.444	1.000	NS
	M2_13	16	16	0	1.000	**< 0.001**	**	-	-	-	-	
	M2_3	21	20	1	0.952	**< 0.001**	**	13	8	0.381	0.383	NS
	T4_5	21	19	2	0.905	**< 0.001**	**	12	9	0.429	1.000	NS
LFB_16	M3_10	17	17	0	1.000	**< 0.001**	**	9	8	0.471	1.000	NS
	M2_2	16	16	0	1.000	**< 0.001**	**	-	-	-	-	
	M3_19	10	10	0	1.000	**0.002**	**	6	4	0.400	0.754	NS
	M2_19	10	-	-	-	-		7	3	0.700	0.344	NS
	M2_3	10	-	-	-	-		6	4	0.400	0.754	NS
LFB_19	M3_19	10	-	-	-	-		5	5	0.500	1.000	NS
	M2_13	10	-	-	-	-		5	5	0.500	1.000	NS
LFB_24	M3_19	-	-	-	-	-		5	3	0.625	0.727	NS
	T2_6	10	10	0	1.000	**0.002**	**	-	-	-	-	
	T4_5	10	10	0	1.000	**0.002**	**	-	-	-	-	

Appendix 5

Supplementary information for Chapter 5

Supplementary Table S1. Raw and filtered ASE sites found in F1 males.

ASE sites	**Between replicates**	**Shared by all replicates**	**Final set (after SNP filtering)**
Total	10,546	5,757	605
Exonic	6,341	4,351	450
Intronic	922	272	49
Intergenic	3,208	1,072	104
Orphan	75	32	2

Supplementary Table S2. Summary statistics of read depth and allele expression bias p_B. Mean, standard deviation and median for each annotation feature are indicated.

ASE sites	**Read depth**						**Allele expression bias p_B**					
Tissue	**BH**			**HB**			**BH**			**HB**		
	Mean	*SD*	*Med*	*Mean*	*SD*	*Med*	*Mean*	*SD*	*Med*	*Mean*	*SD*	*Med*
Exonic	119.9	230.5	56.0	116.3	242.9	49.5	0.62	0.13	0.62	0.48	0.16	0.49
Intronic	81.3	59.6	56.5	99.0	80.3	65.5	0.55	0.15	0.51	0.47	0.15	0.47
Intergenic	47.8	44.5	34.7	43.3	38.0	33.1	0.57	0.21	0.59	0.46	0.23	0.46
Orphan	56.5	19.8	56.5	50.1	18.5	50.1	1	0	1	0	0	0

Supplementary Table S3. Genomic location of genes exhibiting parent-of-origin expression. NS, not significant allelic expression difference between BH and HB. S (H), significant difference and homogeneity across reciprocal cross replicates. S (NH), significant difference and heterogeneity across replicates. Genes with parent-of-origin-specific expression are underlined.

Contig	Gene	Position	p_{BH}	p_{HB}	Category
NW_002987035	PHUM060620	40458-46226	0.61	0.36	S (NH)
	PHUM060670	83559-84713	0.65	0.66	NS
	PHUM060680	*85225-88543*	*0.68*	*0.24*	*S (NH)*
	PHUM060940	104264-105551	0.51	0.34	NS
	PHUM061080	119531-120891	0.62	0.43	S (H)
	PHUM061110	127663-132056	0.64	0.46	S (NH)
	PHUM061240	136229-138956	0.50	0.51	NS
NW_002987337	*PHUM328760*	*430872-436687*	*0.71*	*0.30*	*S (H)*
	PHUM328870	509417-509907	0.48	0.49	NS
	PHUM328880	524255-525120	0.48	0.44	NS
NW_002987464	*PHUM395570*	*249199-250663*	*0.00*	*1.00*	*S (H)*
NW_002987772	PHUM433120	113979-126152	0.52	0.39	NS
	PHUM433160	173723-179777	0.54	0.46	NS
	PHUM433630	451438-454102	0.44	0.61	S (H)
	PHUM433670	462599-471545	0.65	0.44	S (NH)
	PHUM433690	*478833-482035*	*0.83*	*0.27*	*S (H)*
	PHUM433820	490596-498258	0.66	0.52	NS
NW_002987850	PHUM514030	1089249-1090304	0.57	0.20	S (NH)
	PHUM514040	*1090608-1093673*	*0.69*	*0.19*	*S (H)*
	PHUM514070	1099188-1100311	0.55	0.58	NS
	PHUM514080	1101610-1102870	0.55	0.47	NS
	PHUM514100	1108297-1108727	0.53	0.64	NS
	PHUM514160	1146108-1148799	0.64	0.31	S (H)
	PHUM514810	1218544-1221219	0.57	0.52	NS
	PHUM514820	1221826-1223172	0.60	0.45	S (H)
NW_002987860	PHUM544450	15819-16989	0.54	0.43	S (H)
	PHUM544460	18778-19757	0.59	0.53	NS
	PHUM544500	*31966-35476*	*0.66*	*0.33*	*S (NH)*
	PHUM544510	37283-38885	0.73	0.47	S (NH)
	PHUM544520	39872-42107	0.67	0.37	S (NH)
	PHUM544530	42378-43609	0.62	0.30	S (NH)
	PHUM544540	43966-44856	0.51	0.43	S (NH)
	PHUM544550	45252-46440	0.71	0.64	NS
	PHUM544690	54483-61838	0.43	0.43	NS
	PHUM544800	63902-65612	0.62	0.45	S (NH)

	PHUM544810	66021-69079	0.73	0.48	S (NH)
	PHUM545140	101813-103717	0.72	0.39	S (NH)
	PHUM545350	106859-107487	0.58	0.50	S (H)
	PHUM545460	110940-123049	0.56	0.49	S (NH)
	PHUM545470	*113638-114130*	*1.00*	*0.00*	*S (H)*
	PHUM545480	*114685-115183*	*1.00*	*0.00*	*S (NH)*
	PHUM545490	*116229-120441*	*0.73*	*0.33*	*S (NH)*
	PHUM545510	125652-127929	0.69	0.39	S (NH)
NW_002987863	PHUM549150	22136-24285	0.51	0.52	NS
	PHUM549260	28255-30813	0.53	0.53	NS
	PHUM549380	35761-40862	0.66	0.45	S (H)
	PHUM549390	41588-42109	0.60	0.76	NS
	PHUM549400	*42268-43113*	*0.67*	*0.33*	*S (H)*
	PHUM549550	83274-84538	0.50	0.59	NS
	PHUM549570	121859-123904	0.47	0.44	NS
	PHUM549590	131624-133767	0.55	0.53	NS
	PHUM551020	369134-371226	0.60	0.59	NS
	PHUM551030	371654-374605	0.61	0.57	NS
NW_002987884	PHUM595770	28334-29415	0.70	0.36	S (H)
	PHUM595780	41214-44477	0.53	0.46	NS
	PHUM595810	61658-61976	0.43	0.65	S (H)
	PHUM596220	195331-213539	0.56	0.42	S (H)
	PHUM596750	*630433-641590*	*0.16*	*0.71*	*S (H)*
	PHUM596770	650003-660984	0.14	0.27	NS
NW_002988572	*PHUM622880*	*232-1127*	*1.00*	*0.00*	*S (H)*

Supplementary Table S4. Top 10 significant genes (FDR-adjusted posterior probability of equal expression < 0.05) with highest fold expression change between pure body (BB) and head (HH) louse males.

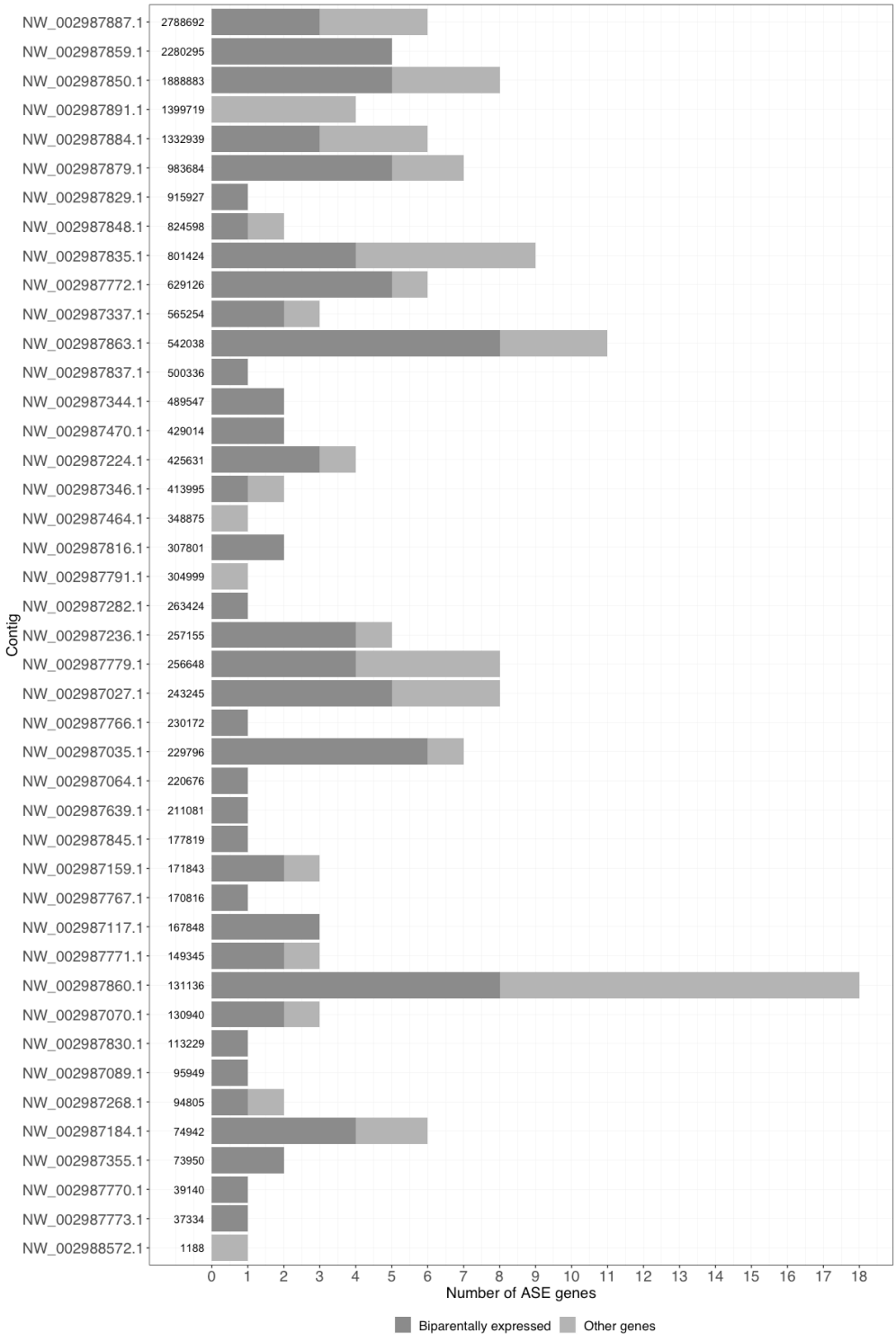
Gene	Enriched	Log_2 fold change	PPEE	Function
PHUM619270	BB	8.05	0.000000	Chloride intracellular channel
PHUM420990	BB	8.00	0.000007	Hypothetical protein
PHUM492240	BB	6.81	0.000988	Cuticle protein
PHUM478280	BB	6.67	0.000247	Hypothetical protein
PHUM098910	BB	5.55	0.000000	Cytochrome P-450
PHUM058910	BB	4.30	0.033027	Cuticle protein
PHUM192050	BB	3.04	0.000000	Conserved Hypothetical protein
PHUM156990	BB	2.21	0.004483	Amino acid transporter
PHUM456970	BB	1.96	0.002021	Harpin hrpN
PHUM595820	BB	1.94	0.004526	Protein takeout precursor
PHUM623280	HH	-6.74	0.000000	Conserved hypothetical protein
PHUM586700	HH	-4.66	0.000000	Cytochrome b5 isoform
PHUM365700	HH	-4.17	0.000000	Defensin precursor
PHUM494820	HH	-3.01	0.000300	Hemocyanin subunit
PHUM242850	HH	-2.54	0.000000	Conserved Hypothetical protein
PHUM087840	HH	-2.23	0.000062	Cytochrome P450
PHUM233900	HH	-2.04	0.000004	Parathyroid hormone receptor
PHUM358590	HH	-1.98	0.000000	Hypothetical protein
PHUM216730	HH	-1.93	0.000000	Vitamin k-dependent gamma-carboxylase
PHUM058890	HH	-1.84	0.007946	Adult-specific cuticular protein ACP-20 precursor

Supplementary Table S5**.** All 74 genes with significant expression difference (FDR-adjusted posterior probability of equal expression < 0.05) between BH and HB males. The corresponding differential expression pattern in pure males is indicated.

Gene	Enriched	Log_2 fold change	PPEE	In pure males	Function
PHUM146180	BH	8.27	0.000001	=	Hypothetical protein
PHUM221760	BH	7.73	0.000003	=	Hypothetical protein
PHUM146230	BH	7.68	0.000004	=	Hypothetical protein
PHUM460950	BH	7.66	0.000007	=	Spectrin beta chain
PHUM403440	BH	7.55	0.000012	=	Conserved hypothetical protein
PHUM075010	BH	7.40	0.000016	=	Conserved hypothetical protein
PHUM534950	BH	7.36	0.000013	=	Protein takeout precursor
PHUM595880	BH	7.26	0.000073	=	Uncharacterised
PHUM460960	BH	7.18	0.000034	=	Hypothetical protein
PHUM596000	BH	7.12	0.000154	=	Uncharacterised
PHUM146120	BH	7.11	0.000037	=	Hypothetical protein
PHUM064140	BH	7.10	0.000045	=	Hypothetical protein
PHUM146140	BH	6.68	0.000114	=	Hypothetical protein
PHUM440030	BH	6.57	0.000124	=	P protein
PHUM127900	BH	6.47	0.000272	Not expr.	Ves G1 allergen precursor
PHUM492030	BH	6.46	0.000211	=	Conserved hypothetical protein
PHUM594660	BH	6.23	0.000550	=	Ejaculatory bulb-specific protein 3 precursor
PHUM146600	BH	6.16	0.000451	=	Uncharacterised
PHUM463140	BH	5.99	0.000973	Not expr.	Hypothetical protein
PHUM595890	BH	5.96	0.001050	=	Hypothetical protein
PHUM038270	BH	5.95	0.001443	=	Vitellogenin receptor
PHUM410500	BH	5.75	0.002811	=	Transcription factor MafG
PHUM207320	BH	5.64	0.002163	=	APC/C activator protein CDH1
PHUM116840	BH	5.57	0.002932	=	Conserved hypothetical protein

PHUM410490	BH	5.47	0.008063	=	Hypothetical protein
PHUM600930	BH	5.29	0.013729	=	Hypothetical protein
PHUM534910	BH	5.05	0.011973	=	Chromaffin granule amine transporter
PHUM146130	BH	4.77	0.031859	=	Hypothetical protein
PHUM448870	BH	4.64	0.036916	=	Glucose dehydrogenase precursor
PHUM390430	BH	3.76	0.032225	=	Lachesin precursor
PHUM365700	BH	1.95	0.000000	H	Uncharacterised
PHUM419870	BH	1.75	0.000335	=	Alcohol dehydrogenase
PHUM581710	BH	1.56	0.000000	=	Synaptonemal complex protein ZIP1
PHUM531990	BH	1.28	0.022552	B	Enzymatic polyprotein
PHUM418740	BH	0.97	0.028497	=	Hypothetical protein
PHUM433660	BH	0.80	0.004527	=	Zinc finger protein RTS2
PHUM474930	BH	0.79	0.012766	=	40 kDa peptidyl-prolyl cis-trans isomerase
PHUM233590	BH	0.53	0.010258	=	Deltex
PHUM490690	BH	0.52	0.000001	=	Predicted protein
PHUM544810	BH	0.46	0.046684	=	Cysteinyl-tRNA synthetase
PHUM574310	BH	0.43	0.003982	=	Conserved hypothetical protein
PHUM216710	BH	0.40	0.000000	=	Hypothetical protein
PHUM255010	BH	0.39	0.032268	=	Conserved hypothetical protein
PHUM449660	BH	0.36	0.028383	=	COMM domain-containing protein
PHUM310380	BH	0.35	0.037982	=	Carbohydrate sulfotransferase
PHUM123610	BH	0.32	0.000467	=	Protein Kr-H2
PHUM617000	BH	0.23	0.010940	=	Conserved hypothetical protein
PHUM268830	BH	0.22	0.000257	=	Cytidine deaminase
PHUM075750	BH	0.20	0.010380	=	Membrin
PHUM233900	HB	-4.88	0.000529	H	Parathyroid hormone receptor
PHUM351580	HB	-3.50	0.019194	=	Hypothetical protein
PHUM173410	HB	-3.45	0.029735	H	Cardioactive peptide precursor
PHUM507350	HB	-1.44	0.048612	=	predicted protein
PHUM149610	HB	-1.41	0.010933	=	Hypothetical protein

PHUM394220	HB	-1.33	0.000144	=	Cysteine desulfurylase
PHUM574160	HB	-1.24	0.000772	=	DNA methyltransferase
PHUM532120	HB	-1.02	0.049589	=	Conserved hypothetical protein
PHUM392580	HB	-1.00	0.000000	=	Brachyury
PHUM129250	HB	-0.94	0.011349	=	Hypothetical protein
PHUM425190	HB	-0.93	0.024275	=	Hypothetical protein
PHUM137310	HB	-0.91	0.015634	=	Conserved hypothetical protein
PHUM602670	HB	-0.79	0.012290	H	Formin 12/cappuccino
PHUM213180	HB	-0.75	0.000000	=	Conserved hypothetical protein
PHUM503110	HB	-0.69	0.016360	=	Conserved hypothetical protein
PHUM348500	HB	-0.61	0.000503	=	Protocadherin-16 precursor
PHUM213170	HB	-0.56	0.000002	=	Homeobox protein extRadenticle
PHUM079970	HB	-0.56	0.040322	=	Hypothetical protein
PHUM228960	HB	-0.51	0.014883	=	Conserved hypothetical protein
PHUM617710	HB	-0.47	0.007150	=	Glutamic acid-rich protein precursor
PHUM500240	HB	-0.46	0.000159	=	WD-repeat protein
PHUM190870	HB	-0.44	0.003067	=	Zfh4
PHUM566690	HB	-0.40	0.005631	=	Protein FAM49B
PHUM392340	HB	-0.31	0.031523	=	Mitochondrial 2-oxodicarboxylate carrier
PHUM125760	HB	-0.26	0.000002	=	Transmembrane and coiled-coil domains protein



Supplementary Figure S1**.** Contig distribution of genes with allele-specific information in the JCVI_LOUSE_1.0 assembly.

Addendum

Published version of Chapter 4

Medical and Veterinary Entomology

Medical and Veterinary Entomology (2018) **32**, 226–234 doi: 10.1111/mve.12287

The unusual reproductive system of head and body lice (*Pediculus humanus*)

A. G. DE LA FILIA[1], S. ANDREWES[2], J. M. CLARK[2] and L. ROSS[1]

[1]School of Biological Sciences, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, U.K. and [2]Departnent of Veterinary and Animal Sciences, University of Massachusetts Amherst, Amherst, MA, U.S.A.

Abstract. Insect reproduction is extremely variable, but the implications of alternative genetic systems are often overlooked in studies on the evolution of insecticide resistance. Both ecotypes of *Pediculus humanus* (Phthiraptera: Pediculidae), the human head and body lice, are human ectoparasites, the control of which is challenged by the recent spread of resistance alleles. The present study conclusively establishes for the first time that both head and body lice reproduce through paternal genome elimination (PGE), an unusual genetic system in which males transmit only their maternally derived chromosomes. Here, we investigate inheritance patterns of parental genomes using a genotyping approach across families of both ecotypes and show that heterozygous males exclusively or preferentially pass on one allele only, whereas females transmit both in a Mendelian fashion. We do however observe occasional transmission of paternal chromosomes through males, representing the first known case of PGE in which whole-genome meiotic drive is incomplete. Finally, we discuss the potential implications of this finding for the evolution of resistance and invite the development of new theoretical models of how this knowledge might contribute to increasing the success of pediculicide-based management schemes.

Key words. Pediculus humanus, human louse, paternal genome elimination, pseudo-haplodiplody, resistance evolution.

Introduction

One of the most striking features of insects is the extraordinary diversity of their reproduction, which is unparalleled in any other animal group. This is illustrated by the wide heterogeneity of reproductive and genetic systems found across insect taxa that differ from the canonical diplodiploidy prevalent in metazoans (Normark, 2003). The most well-known example of these alternative genetic systems is arguably arrhenotoky (i.e. haplodiploidy *sensu stricto*, whereby males develop from unfertilized eggs). However, many other more complex and bizarre non-diplodiploid systems have been described. Many of these are common in economically important insects: for instance, parthenogenesis (female reproduction without fertilization) is disproportionately abundant in pest species, including representatives of groups such as mites, aphids and scale insects, compared with non-pest relatives (Hoffmann *et al.*, 2008; Ross *et al.*, 2013). Telling signs of alternative genetic systems are non-Mendelian inheritance patterns of traits or genetic markers, which are often discovered fortuitously in certain species but are rarely explored further despite their potential implications for key aspects of insect management, such as the evolution of virulence and insecticide resistance.

One of these species is the human louse, *Pediculus humanus*, a blood-sucking ectoparasite that occurs worldwide and causes infestations with serious medical, economic and social consequences. Human lice are divided into two ecotypes: the head louse (*Pediculus humanus capitis*) and the body louse (*Pediculus humanus humanus*) (Durden & Musser, 1994), which differ in their ecology and clinical importance. Whereas body lice feed on human skin and lay eggs on clothes, head lice live and feed on the human scalp only. Epidemiologically, head louse infestations are more common and mostly affect children, regardless of economic status or geographic region (Clark *et al.*, 2013).

Correspondence: Andrés G. de la Filia, School of Biological Sciences, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3FL, U.K. Tel.: + 44 31 650 5464; Fax: + 44 31 650 6564; E-mail: a.gdelafilia@ed.ac.uk

By contrast, body louse infestations are associated with lower socioeconomic conditions and pose a more serious health threat because the body louse is a vector of epidemic pathogenic bacteria, including *Bartonella quintana* (trench fever), *Borrelia recurrentis* (relapsing fever) and *Rickettsia prowazekii* (epidemic typhus) (Raoult & Roux, 1999).

Control of human lice generally involves a combination of manual removal techniques and the use of diverse chemicals often referred to as pediculicides. However, many of the most widely used pediculicides have become ineffective as a result of the spread of resistant strains [see Durand *et al.* (2012) and references therein] and, as many pediculicides share common chemistry and targets (Clark *et al.*, 2013), further spread of resistance is likely. To reduce this risk, it is important to unravel the molecular and metabolic mechanisms involved in pediculicide resistance (Oakeshott *et al.*, 2003), which have been addressed by several studies in recent years (Yoon *et al.*, 2008; Kwon *et al.*, 2014). However, current understanding of how resistance evolves and spreads through populations is very limited because key factors such as population structure, gene flow, reproductive genetics, life history and mating system remain insufficiently explored. Better understanding of these factors and their roles in the evolution of pesticide resistance will support the development of successful novel treatment strategies and management programmes aimed at preventing the spread of resistance genotypes.

Until recently, it was assumed that inheritance of traits such as pesticide resistance in lice would follow the classic laws of Mendelian genetics. However, an unexpected finding in the body louse suggested that whereas allele transmission in females followed Mendelian expectations, it was non-Mendelian in males: heterozygous male parents systematically passed on one of their two alleles to their offspring (McMeniman & Barker, 2005). Moreover, the transmitted allele was of maternal origin in all cases and the paternally derived alternative allele was absent from the offspring. This mode of inheritance is consistent with paternal genome elimination (PGE), a type of haplodiploid reproduction found across several arthropod orders in which males do not transmit paternally inherited alleles to their offspring (Normark, 2003). It is surprising that the possible presence of PGE in *P. humanus* has not been considered in the context of louse control because it may have implications for the evolution of pesticide resistance. Theoretical approaches have shown that haplodiploidy can accelerate the invasion of resistant alleles under certain circumstances (Crozier, 1985; Caprio & Hoy, 1995; Denholm *et al.*, 1998; Carrière, 2003), and PGE has been invoked to explain the rapid spread of insecticide resistance in New Caledonian populations of the coffee berry borer beetle *Hypothenemus hampei* (Brun *et al.*, 1995). Furthermore, PGE is likely to elicit sex-specific responses and selection pressures that can further affect the way resistance genotypes spread through populations (Carrière, 2003).

Although the study by McMeniman & Barker (2005) is suggestive of the presence of PGE in *P. humanus*, it requires further confirmation. They show that a proportion of heterozygous males transmit both alleles in a Mendelian fashion, which would mean that PGE was polymorphic in the study population (McMeniman & Barker, 2005). This finding is unlike any form of PGE described so far, which has always been found to be complete. Further, McMeniman & Barker (2005) used only three markers in their study, which falls short of covering the whole genome and does not allow determination of whether drive is complete or restricted to some chromosomes. Moreover, the Culpepper strain (Culpepper, 1944) used by McMeniman & Barker (2005) in their experiment might not be representative of natural populations as it has evolved under laboratory conditions since 1945 and has adapted to rabbit blood, rather than human. It is therefore possible that a drive factor emerged in this strain independently of natural body louse populations, which were not sampled. Finally, the study by McMeniman & Barker (2005) included only body lice and no data on inheritance in head lice have been published since. Here, we study patterns of allele inheritance in both head and body louse families derived from recently collected natural populations reared on human blood.

In order to determine whether males show complete genome-wide meiotic drive consistent with PGE, we used a two-generation experimental crossing design and a panel of multiple polymorphic microsatellite markers. Transmission patterns were determined by genotyping both parents and their offspring to determine whether both alleles at a given heterozygous parental locus are present at a 50 : 50 ratio in the offspring (Mendelian transmission) or whether only one allele is transmitted by male parents (PGE). The current study provides the first reported evidence of PGE in the head louse and confirms its occurrence in body lice, albeit with no consistent evidence of a PGE polymorphism between males. We do, however, observe occasional leakage of paternal alleles, especially in body lice. Finally, we also suggest subsequent research directions aimed at increasing current understanding of how PGE operates in lice, particularly whether it affects gene expression patterns in males, and discuss the implications of this unusual genetic system for the evolution of parasitic lice in general and, most specifically, the evolution of pediculicide resistance.

Materials and methods

Experimental design

A series of intraspecific crosses were set up using individuals from the head louse strain SF-HL and the body louse strain Frisco-BL. The SF-HL colony was established in 2002 from head lice collected from ~20 infested children in Plantation, Miami and Homestead (FL, U.S.A.). Approximately 50 males and 50 females were used to temporarily establish a colony on human volunteers (Takano-Lee *et al.*, 2003). Fertile eggs from Homestead were added to the colony at least three times between 2002 and 2006. Approximately 30–50 eggs were introduced each time. The sex ratio of the eggs was assumed to be ~50 : 50. The colony was placed on an *in vitro* rearing system in 2006 (Yoon *et al.*, 2006). The Frisco-BL colony of human body lice was originally collected from nine homeless individuals in San Francisco (CA, U.S.A.) by Dr Jane Koehler (University of California San Francisco Medical Center, San Francisco, CA, U.S.A.) in December 2008. Both colonies have been maintained by the Clark Laboratory at the University of Massachusetts-Amherst on human blood using

the same *in vitro* rearing system (Yoon *et al.*, 2006) under environmental conditions of 30 °C, 70% relative humidity and an LD 16 : 8 h photoperiod in rearing chambers (University of Massachusetts-Amherst Institutional Review Board approval no. E1404/001-002).

Parental generations (F_0) were established by random selection of pairs of sexually immature third instar lice from each colony. These pairs were transferred to individual cages. Lice were sexed after reaching reproductive maturity using the approach first described by Meinking (1999) and cages were checked for same-sex pairs. In these cases, a pair of male-only and female-only cages was selected at random and a randomly chosen individual was swapped between cages. After this point, all cages were screened daily to check for oviposition or the death of parents. Males were removed and stored in 100% ethanol at 4 °C after 7 days or immediately after their death. Females were allowed to lay eggs for 2 weeks and were then removed and stored in 100% ethanol at 4 °C. Offspring (F_1) of all crosses were raised until early third instar stage and then transferred to ethanol. In total, F_1 broods for 26 head and 13 body louse families were obtained.

DNA extraction and polymerase chain reaction

Total genomic DNA from parents and body louse F_1 individuals was extracted with a DNeasy Blood and Tissue Extraction Kit (Qiagen Benelux BV, Venlo, the Netherlands). DNA from head louse F_1 individuals was extracted with a prepGEM Insect Kit (ZyGEM NZ Ltd, Hamilton, New Zealand) in a 20-μL reaction volume. A panel of three multiplexes (MX1, MX2 and MX4) from Ascunce *et al.* (2013) containing 12 microsatellite loci in total (T8_1, M3_10, M3_19, M2_2, T2_6, M2_19, M2_13, M2_3, T9_6, T2_7, T4_5 and T1_4) was used for polymerase chain reaction (PCR) amplification. The PCR reactions for each of the three multiplexes were performed using the Type-it Microsatellite PCR Kit (Qiagen Benelux BV) in a 15-μL reaction volume. Primer sequences and reaction mixes were as described in supplementary Tables S1–3 in Ascunce *et al.* (2013). The PCR reactions were performed under the following conditions: initial denaturation at 95 °C for 5 min; 35 cycles of denaturation at 94 °C for 30 s; annealing at 52 °C for 45 s; extension at 72 °C for 45 s, and a final extension step at 72 °C for 40 min. One microlitre of PCR product for each sample and multiplex was sent to Edinburgh Genomics (University of Edinburgh) for genotyping on the ABI 3730 DNA Analyzer system (ThermoFisher Scientific, Inc., Waltham, MA, U.S.A.).

Primer mapping

To reveal the extent of the genome coverage of the microsatellite panel in use, all loci were mapped against the most recent publicly available louse genome assembly. All forward and reverse primer sequences were blasted against the U.S. Department of Agriculture strain genomic assembly (PhumU2) using the BLAST tool in VectorBase (National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, U.S.A.) with default settings.

Microsatellite scoring and data analysis

Upon reception of raw trace files, microsatellite alleles were scored using the Microsatellite Plugin implemented in GENEIOUS Version 8.1.3 (Biomatters Ltd, Auckland, New Zealand). Estimates of observed (H_O) and expected (H_E) heterozygosity, number of alleles and inbreeding coefficient F_{IS} (Weir & Cockerham, 1984) per locus for F_0 populations were obtained using the online version of GENEPOP Version 4.2 (Raymond & Rousset, 1995; Rousset, 2008) with default parameters. For each family and locus, paternal and maternal allele transmission ratios were calculated as the number of occurrences of one of the two alleles in the F_1 offspring divided by the total number of F_1 genotypes. Because of the clear expectation of allele transmission following McMeniman & Barker (2005) and other PGE species, these parental ratios were defined in different ways to represent these different sex-specific transmission patterns. For paternal transmission ratios, the allele used in this calculation was that with higher representation in the offspring genotypes. For maternal transmission ratios, one of the two alleles was chosen at random. Likewise, when both parents were heterozygous for the same alleles at a given locus, parental allele counts were assigned under the assumption that the driving allele present in all offspring was paternally derived. Exact binomial tests to detect significant deviations from Mendelian expectations in each locus were performed in R Version 3.2.4 (R Foundation for Statistical Computing, Vienna, Austria). To correct for multiple testing, a reduced significance level of $\alpha = 0.01$ is considered in addition to the conventional level of $\alpha = 0.05$.

Results

Informative parents and microsatellite panel

In order to determine patterns of allele transmission, the F_1 offspring of F_0 parents with at least one heterozygous locus were genotyped because parents that are homozygous for all loci are not informative. Multi-locus heterozygosity of parental populations was higher in head louse F_0 ($H_O = 0.351$) than in body louse F_0 ($H_O = 0.256$) despite higher allelic richness in the latter, as a result of the elevated inbreeding coefficient in the body louse population ($F_{IS} > 0.2$) (Table 1). At least one heterozygous marker was found in all 26 head louse and 11 body louse fathers. Likewise, 24 head louse and all 13 body louse mothers were heterozygous for at least one locus. This allowed for the determination of both paternal and maternal allele transmission patterns in almost all families (Table S1, online).

All F_0 and F_1 individuals were genotyped using the 12-locus microsatellite panel, but not all markers could be included in the analysis (Tables S2 and S3, online). T9_6 was monomorphic in head lice, whereas T9_6 and T1_4 failed to amplify in most body louse individuals and were excluded in this ecotype, but all remaining loci were polymorphic and amplified consistently in most families. It was further decided that the T8_1 locus should be excluded in both ecotypes as a result of genotype inconsistencies. Therefore, from the initial 12-locus microsatellite

Table 1. Multi-locus descriptive statistics of head and body louse F_0 parental populations.

Species	Families	F_1/family	Loci	Allele/locus	H_O	H_E	F_{IS}	H_O ♂	H_O ♀
Head louse	26	8–12	11 (10)	2.55 ± 0.32	0.341 ± 0.065	0.366 ± 0.071	0.021 ± 0.044	0.315 ± 0.064	0.367 ± 0.069
Body louse	13	7–22	9 (9)	3.00 ± 0.21	0.256 ± 0.056	0.336 ± 0.051	0.262 ± 0.101	0.239 ± 0.064	0.274 ± 0.067

Families, number of F_0 parental pairs establishing F_1 broods.
Loci, number of reliable loci included in the analysis (informative; i.e. polymorphic loci in parentheses).
F_1/family, range of number of individuals per family genotyped for each ecotype.
Allele/locus, mean ± standard error (SE) number of alleles per marker.
H_O, mean ± SE observed heterozygosity across all loci.
H_E, mean ± SE expected heterozygosity across all loci.
F_{IS}, mean ± SE F_{IS} across all loci (following Weir & Cockerham, 1984).
H_O ♂ and H_O ♀, mean ± SE observed heterozygosity across all loci for F_0 fathers and F_0 mothers.

Table 2. Genome location of markers developed by Ascunce *et al.* (2013) (used in this study) and Leo *et al.* (2002) [used in McMeniman & Barker (2005)].

Panel	Locus	Scaffold	*E*-value
Ascunce *et al.* (2013)	M3_10	DS235157	0.002
	M3_19	DS235833	0.002
	M2_2	DS235048	0.0005
	T2_6	DS235090	0.0002
	M2_19	DS235875	0.0005
	M2_13	DS235785	0.002
	M2_3	DS235111	0.0005
	T9_6	DS235882	< 0.0001
	T2_7	DS235100	0.002
	T4_5	DS235283	0.0002
	T1_4	DS235023	0.002
Leo *et al.* (2002)	ML_8	DS235886	< 0.0001
	ML_9	DS235886	0.002
	ML_10	DS235042*	0.023
		DS235005†	0.98

All forward and reverse primers for each locus mapped to the same scaffold; the highest *E*-value for each of the pairs is shown, except for ML_10 (*forward; †reverse).

panel, 10 and nine reliable informative loci were used to estimate allele transmission patterns in head and body louse families, respectively.

To assess the genome coverage of the microsatellite panel, all primer sequences were blasted to the *P. humanus* genome assembly to determine whether they were located in different genomic regions. Each of the markers was found to map to a distinct scaffold in the genome assembly (Table 2). Although the genome assembly does not allow for the exact determination of which chromosomes are targeted by the markers used herein, the present authors are confident that the panel offers sufficient coverage for a genomewide study of transmission patterns. By contrast, very limited genome coverage of the three markers used in McMeniman & Barker (2005) was found because two of them map to the same scaffold and the location of the third is unclear.

Allele transmission patterns in males and females

For most families and loci in both ecotypes, heterozygous head and body louse males did not transmit alleles in a Mendelian fashion, but consistently passed on only one allele to the F_1. By contrast, females transmit both alleles to their offspring (Fig. 1, Table S1). These patterns are consistent with PGE: females are normally diploid and exhibit Mendelian transmission, whereas males show whole-genome drive and transmit only the maternally inherited allele at each locus.

However, despite clear preferential transmission of one of the two alleles at each locus, head and body louse males sporadically also transmitted alternative (i.e. paternally inherited) alleles. Occasional paternal transmission of alternative alleles was observed across most markers, except for M2_19, M3_19 and M3_10 (Fig. 2). Escapes were rare in head louse males: three males (LFH_20, LFH_33 and LFH_34) passed on an alternative allele once at a single different locus (T1_4, M2_2 and T2_7, respectively). The other 23 head louse males showed complete PGE at all heterozygous loci. Overall, 64 of 71 head louse paternal transmission ratios deviated significantly from the Mendelian expectation of equal transmission at a significance threshold of 0.01 (all 71 at $\alpha = 0.05$), compared with one of 81 ratios in head louse females.

In body louse families, incomplete PGE occurrences were more frequent. Four of the 11 informative males also transmitted the alternative allele at least once (LBH_01 and LBH_09 at one locus, LBH_06 and LBH_15 at two loci). With a significance threshold of 0.01, 19 of 28 ratios deviated from Mendelian transmission (25 of 28 at $\alpha = 0.05$). In body louse females, none of the 29 transmission ratios deviated from Mendelian expectations.

The present study did not find a consistent pattern of incomplete PGE instances across families and loci. To exclude genotyping error for these unexpected paternal escapes, both parents and offspring were re-genotyped and additional offspring were genotyped when available. We are therefore confident that the current findings represent true events of paternal chromosomes escaping germline elimination at low frequencies, particularly in body lice.

Discussion

The allele transmission patterns described in the present study offer conclusive evidence of a genome-wide male transmission ratio distortion in both ecotypes of *P. humanus*: males exclusively (or, in some cases, preferentially) transmit only one

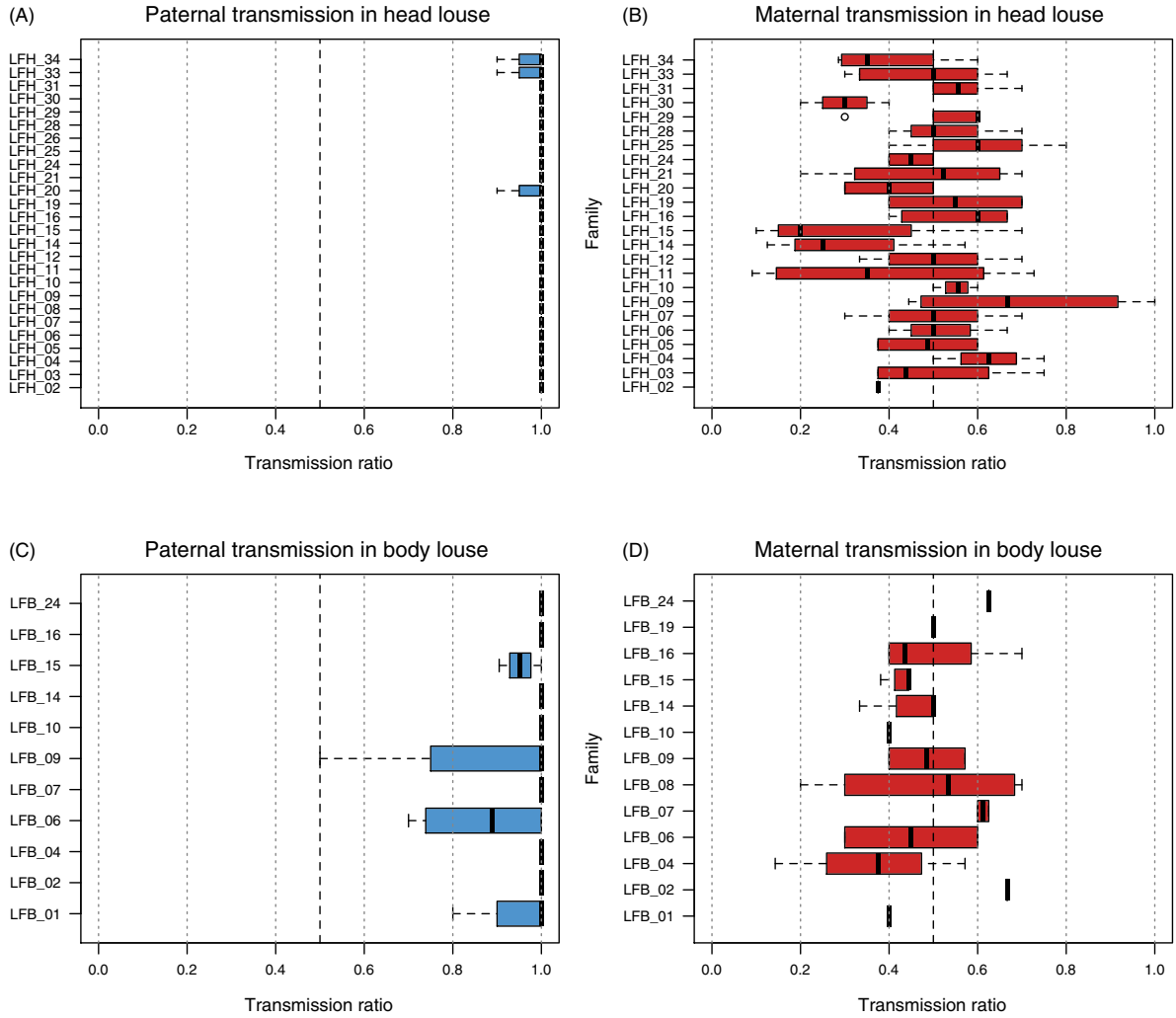


Fig. 1. Allele transmission ratios across all loci for head and body louse (A, C) males and (B, D) females. When both alleles are equally represented in F_1 offspring, the transmission ratio is 0.5 (complete Mendelian transmission). A transmission ratio of 1 indicates complete drive of one of the parental alleles. [Colour figure can be viewed at wileyonlinelibrary.com].

of their alleles to their offspring. In addition, heterozygous genotypes in males from both ecotypes unambiguously indicate that males are diploid and that both paternally and maternally inherited chromosomes are kept in the soma. Although the two-generation experimental design used in this study does not explicitly allow for determination of the parental origin of alleles in F_0 individuals, McMeniman & Barker (2005) already demonstrated that driving alleles were always maternally inherited in body louse males. All these findings are consistent with germline PGE, a pseudohaplodiploid genetic system in which males develop from fertilized eggs and are diploid, but eliminate chromosomes of paternal origin from their germline. This type of reproduction is also found in several other insect taxa such as mealybugs, the coffee borer beetle, two dipteran clades and book lice (Burt & Trivers, 2006; Gardner & Ross, 2014; de la Filia *et al.*, 2015; Hodson *et al.*, 2017).

All males in the present study exhibited whole-genome transmission ratio distortion with sporadic, inconsistent leakages of non-driving alleles in some individuals. Interestingly, the current data reveal that leakages are more frequent in body than in head lice, although the power to detect these occurrences was greater in the latter because twice as many head louse families were screened and they showed higher levels of heterozygosity. However, the study found no evidence of a female PGE-inducing genetic polymorphism as suggested by McMeniman & Barker (2005). In their model, a codominant maternally transmitted genetic element is responsible for elimination of paternal alleles in male offspring so that females that are heterozygous for this element produce PGE sons that pass on only maternal alleles and non-PGE sons that transmit parental alleles in a Mendelian fashion. However, the mapping of markers to the louse genome revealed that McMeniman & Barker (2005) appear to have targeted a single chromosome only. Therefore, an alternative interpretation of these earlier results that is consistent with the sporadic leakage of paternal alleles observed in the current study would be a germline PGE mechanism in which discrimination

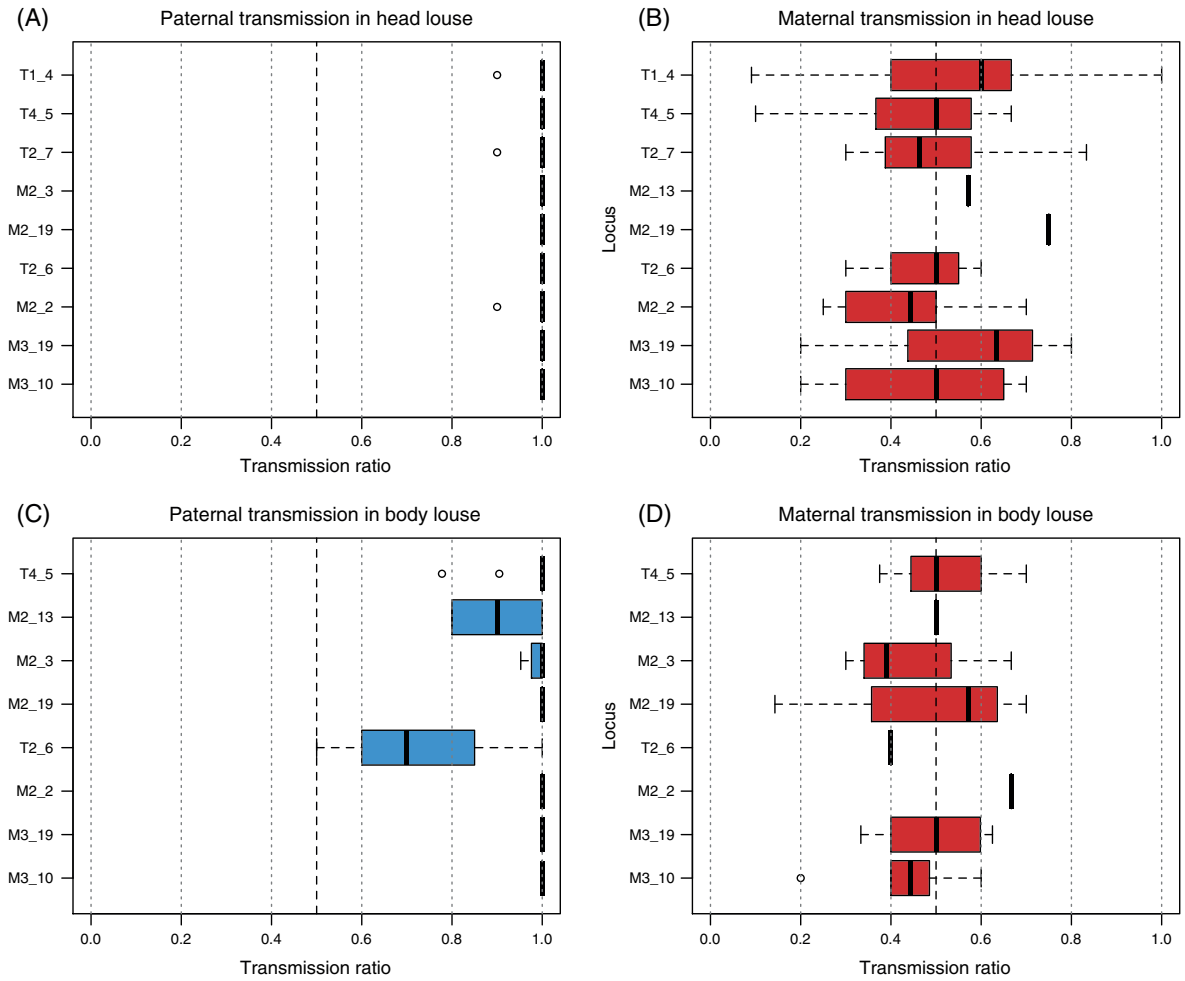


Fig. 2. Paternal and maternal allele transmission ratios in all (A, B) head and (C, D) body louse families grouped by loci. [Colour figure can be viewed at wileyonlinelibrary.com].

between maternal and paternal chromosomes during spermatogenesis is not infallible. In germline PGE, males are somatically diploid and elimination of paternal chromosomes is achieved via non-random assortment of chromosomes during meiosis so that only nuclei containing maternal chromosomes develop into viable sperm (Burt & Trivers, 2006). Whole-genome meiotic drive of maternal chromosomes in germline PGE taxa has been most extensively described in sciarid flies (Esteban *et al.*, 1997; Goday & Esteban, 2001) and mealybugs (Bongiorni *et al.*, 2004, 2009). Allele transmission patterns in louse males reveal that paternal chromosomes are similarly excluded from active spermatocytes, but are occasionally able to escape elimination by migrating with other maternal chromosomes in lieu of their homologues, particularly in body lice. Achiasmatic male meiosis, which is an imperative requisite for PGE as it prevents mixing of paternal and maternal alleles, has been documented in lice (Tombesi & Papeschi, 1993; Tombesi *et al.*, 1999; Bressa *et al.*, 2015). As recombination between maternal and paternal homologues cannot account for transmission of paternal alleles, the detected leakage would encompass entire paternal chromosomes. Therefore, the apparent non-PGE body louse males found by McMeniman & Barker (2005) are more likely to be males exhibiting biparental transmission for the chromosome targeted by their marker panel only, whereas transmission of other chromosomes consistent with PGE (or additional occurrences of paternal leakages) would have passed undetected.

Head and most particularly body lice are the first species for which incomplete (albeit not polymorphic) PGE has been explicitly reported. The discrimination mechanism by which paternal and maternal louse chromosomes are differentially tagged is unknown. In other PGE taxa, maternal and paternal chromosomes differ in patterns of DNA methylation (Bongiorni *et al.*, 1999, 2009) and histone modifications (Goday & Ruiz, 2002; Greciano & Goday, 2006; Khosla *et al.*, 2006; Escribá *et al.*, 2011; Prantera & Bongiorni, 2012), which may mediate discrimination between homologues during spermatogenesis. In lice, inaccuracies of the parent-of-origin discrimination mechanism, whichever its nature, could result in the occasional migration of paternal chromosomes with the remaining maternal chromosomes.

Although at this stage the issue of how these leakages occur is subject only to speculation, a likely PGE mechanism in

which only nuclei containing maternal chromosomes develop into viable sperm (bar accidental leakage of paternal homologues) can be proposed based on previous cytogenetic work in lice. Louse spermatogenesis is highly complex: achiasmatic meiosis is followed by three or four mitotic divisions to yield a 32/64-cell cyst that undergoes a final and unequal mitosis in which most cytoplasmic material is excluded from half the cells, which degenerate into pyknotic nuclei (Hindle & Pontecorvo, 1942; Bressa *et al.*, 2015) similar to those seen in mealybug spermatogenesis (Bongiorni *et al.*, 2004, 2009). The present authors agree with McMeniman & Barker (2005) that non-random assortment of chromosomes is likely to occur in the last, unequal division, after which only the spermatids carrying maternal chromosomes develop into viable spermatozoa. This implies an inverted meiotic sequence in which the first division is equational rather than reductional, with sister chromatids separating before segregation of homologous chromosomes, as found in other PGE taxa such as mealybugs (Viera *et al.*, 2008). It is possible that inverted meiosis in louse males has been historically overlooked in cytogenetic studies as a result of the lack of heteromorphic bivalents and the tight association and highly condensed nature of louse chromosomes, which are holocentric [i.e. they lack a localized centromere; see Bressa *et al.* (2015) and references therein]. Recently, Bressa *et al.* (2015) reported that sister chromatid separation indeed occurs in the first division, but conclusive evidence has yet to be presented.

PGE may have important implications for the transmission of pesticide resistance, which must be parent-of-origin-dependent in males. Resistant males are unable to pass on the trait to their offspring when it is paternally derived and hence resistance will be lost through the paternal line even if it is under strong positive selection. By contrast, males that inherited the resistance trait from their mothers will transmit it to all their offspring, rather than half as occurs in Mendelian inheritance. These characteristic PGE inheritance patterns complicate predictions of resistance invasion without models that explicitly consider sex-specific differences on allelic transmission. In addition, PGE also reduces effective population sizes (Wright, 1933), although this effect may be small when sex ratios are female-biased (Hedrick & Parker, 1997), as is often the case in louse populations (Perotti *et al.*, 2004).

Another way in which PGE can affect the evolution of resistance is through its potential effect on patterns of gene expression. Taxa in which PGE occurs vary in the degree of paternal genome expression in males, which can affect response to insecticides and have an impact on rates of resistance evolution. In many PGE groups, paternal chromosomes are lost (haploid soma PGE) or transcriptionally inactive (functionally haploid PGE) (Normark, 2003). One immediate consequence of these two forms of PGE is that maternally inherited recessive alleles are directly exposed to selection in males, as under arrhenotoky. Therefore, the evolution of insecticide resistance is faster in arrhenotokous (Crozier, 1985; Havron *et al.*, 1987; Caprio & Hoy, 1995; Denholm *et al.*, 1998) and functionally haploid PGE species (Brun *et al.*, 1995) than in diplodiploids [but not always; see Carrière (2003)]. However, males in other PGE taxa are diploid and may express both alleles regardless of parental origin (diploid soma PGE) (Gardner & Ross, 2014).

Because of this variation in gene expression patterns in PGE systems, it is important to precisely determine the degree of paternal genome expression in louse males. Although heterozygous males show that paternal chromosomes are retained, it is still possible that these are transcriptionally inert. In functionally haploid PGE taxa that remain somatically diploid, inactive paternal chromosomes appear as highly compacted dots (Brown & Nur, 1972; Brun *et al.*, 1995; Hodson *et al.*, 2017). To the present authors' knowledge, this conspicuous chromosomal behaviour has never been described in human lice, which suggests that PGE is of the diploid soma form and paternal chromosomes are hence transcriptionally active. Phenotypic assays in hybrid individuals are other indicators of paternal chromosome expression in PGE males because they are expected to show the same traits as males from the maternal species if paternal chromosomes are inactivated. Body size and tibia length measurements in hybrids have been reported to be intermediate between head and body lice (Busvine, 1978), but this study did not discriminate between male and female offspring.

If paternal chromosomes are expressed in human louse males, the aforementioned theoretical models on evolution resistance in haplodiploids cannot be applied because they do not consider diploid expression in PGE species with arrhenotokous-like inheritance. Therefore, new theory must be developed to predict how whole-genome meiotic drive in males with diploid gene expression will affect resistance evolution.

How PGE evolved in the human louse remains an open question. Although *P. humanus* is the only anopluran (i.e. sucking louse) in which the occurrence of PGE has been explicitly demonstrated, the same modified spermatogenesis has been reported in other parasitic louse species. These include another anopluran, the pig louse *Haematopinus suis* (Phthiraptera: Haematopinidae) (Bayreuther, 1955; Tombesi & Papeschi, 1993), and members of two suborders of the paraphyletic group Mallophaga (i.e. chewing lice): Amblycera [the guinea pig louse *Gyropus ovalis* (Phthiraptera: Gyropidae) and the chicken body louse *Menacanthus stramineus* (Phthiraptera: Menoponidae)] (Scholl, 1955; Tombesi & Papeschi, 1993) and Ischnocera [two species of *Bovicola* (Phthiraptera: Trichodectidae), the goat louse] (Tombesi *et al.*, 1999). More tellingly, empirical evidence of PGE in a close relative of parasitic lice, the book louse *Liposcelis* sp. (Psocoptera: Liposcelididae), has been recently provided (Hodson *et al.*, 2017). In this species, PGE is of the functionally haploid type, with males retaining condensed paternal chromosomes. Although the phylogenetic relationships between and within Psocoptera (book lice) and Phthiraptera are not yet fully resolved and this division has been called into question (Yoshizawa & Johnson, 2010; Li *et al.*, 2015), there is consensus that all lice form a monophyletic group and it is therefore possible that PGE is common to all of them. Formal investigation of transmission patterns and somatic heterochromatinization in these or other parasitic louse species would be necessary to corroborate this hypothesis.

Several authors have suggested that PGE may have evolved through attempts by endosymbionts to manipulate their host's reproduction (Normark, 2004; Kuijper & Pen, 2010; Ross *et al.*,

2012). The rationale for this is that maternally transmitted endosymbionts benefit from a female-biased sex ratio and that the elimination of paternally derived chromosomes in males may be a way of killing male offspring. Lice harbour several maternally inherited endosymbiotic bacteria including both obligate nutritional mutualists as well as bacteria known to manipulate host reproduction in their own favour, such as *Wolbachia*. Hence, could PGE in lice be induced by endosymbionts? Probably not: human louse males remain diploid throughout development and only eliminate their paternally derived genome during spermatogenesis, which is unlikely to induce male-specific mortality and is therefore not in the interest of the endosymbionts.

The present study demonstrates that PGE is present in both *P. humanus* ecotypes and outlines some considerations of the impact of the particular genetic system on the evolution of pediculicide resistance. A more complete understanding of human louse biology is imperative to facilitate the design and application of successful resistance management programmes. Yet asymmetry in gene transmission patterns, sex ratio bias and possible phenotypic consequences of PGE have not been considered thus far. The characterization and compact nature of the *P. humanus* genome enable genome-wide allele-specific expression studies to determine the extent to which paternally inherited alleles can confer resistance phenotypes in males. If they can, theoretical models of resistance evolution combining diploid expression and haplodiploid transmission will be needed to maximize the success of resistance control strategies.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: DOI: 10.1111/mve.12287

Table S1. Paternal and maternal transmission ratios for all families and loci.

Table S2. Head louse genotypes.

Table S3. Body louse genotypes.

Acknowledgements

This study benefitted from travel funding provided by the National Evolutionary Synthesis Center (NESCent; Durham, NC, U.S.A.). The authors would like to thank David Haig, University of Harvard, for bringing the unusual reproduction of lice to their attention and inspiring this study, and Benjamin Normark, University of Massachusetts Amherst, and Stephen Barker, University of Queensland, for useful comments and discussions. AGdlF was funded by the Darwin Trust of Edinburgh. JMC was funded by a National Institutes of Health, National Institute of Allergy and Infectious Diseases grant for the 'Functional and Transcriptional Profiling of Lice Defenses against Pediculicides' (NIH/NIAID 1 R56AI081933-01A2). LR was funded by the Natural Environment Research Council (Ne/K009516/1).

References

Ascunce, M.S., Toups, M.A., Kassu, G., Fane, J., Scholl, K. & Reed, D.L. (2013) Nuclear genetic diversity in human lice (*Pediculus humanus*) reveals continental differences and high inbreeding among worldwide populations. *PLoS One*, **8**, e57619.

Bayreuther, K. (1955) Holokinetische Chromosomen bei *Haematopinus suis* (Anoplura, Haematopinidae). *Chromosoma*, **7**, 260–270.

Bongiorni, S., Cintio, O. & Prantera, G. (1999) The relationship between DNA methylation and chromosome imprinting in the coccid *Planococcus citri*. *Genetics*, **151**, 1471–1478.

Bongiorni, S., Fiorenzo, P., Pippoletti, D. & Prantera, G. (2004) Inverted meiosis and meiotic drive in mealybugs. *Chromosoma*, **112**, 331–341.

Bongiorni, S., Pugnali, M., Volpi, S., Bizzaro, D., Singh, P.B. & Prantera, G. (2009) Epigenetic marks for chromosome imprinting during spermatogenesis in coccids. *Chromosoma*, **118**, 501–512.

Bressa, M.J., Papeschi, A.G. & Toloza, A.C. (2015) Cytogenetic features of human head and body lice (Phthiraptera: Pediculidae). *Journal of Medical Entomology*, **52**, 918–924.

Brown, S.W. & Nur, U. (1972) Heterochromatic chromosomes in the coccids. *Invertebrate Oogenesis*, **2**, 97.

Brun, L.O., Stuart, J., Gaudichon, V., Aronstein, K. & French-Constant, R.H. (1995) Functional haplodiploidy: a mechanism for the spread of insecticide resistance in an important international insect pest. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 9861–9865.

Burt, A. & Trivers, R. (2006) *Genes in Conflict: The Biology of Selfish Genetic Elements*. Harvard University Press, Cambridge, MA.

Busvine, J.R. (1978) Evidence from double infestations for the specific status of human head lice and body lice (Anoplura). *Systematic Entomology*, **3**, 1–8.

Caprio, M.A. & Hoy, M.A. (1995) Premating isolation in a simulation model generates frequency-dependent selection and alters establishment rates of resistant natural enemies. *Journal of Economic Entomology*, **88**, 205–212.

Carrière, Y. (2003) Haplodiploidy, sex, and the evolution of pesticide resistance. *Journal of Economic Entomology*, **96**, 1626–1640.

Clark, J.M., Yoon, K.S., Lee, S.H. & Pittendrigh, B.R. (2013) Human lice: past, present and future control. *Pesticide Biochemistry and Physiology*, **106**, 162–171.

Crozier, R.H. (1985) Adaptive consequences of male-haploidy. *Spider Mites: Their Biology, Natural Enemies and Control*, Vol. 1A (ed. by W. Helle & M.W. Sabelis), pp. 201–222. Elsevier, Amsterdam.

Culpepper, G.H. (1944) The rearing and maintenance of a laboratory colony of the body louse. *American Journal of Tropical Medicine*, **24**, 327–329.

Denholm, I., Cahill, M., Dennehy, T.J. & Horowitz, A.R. (1998) Challenges with managing insecticide resistance in agricultural pests, exemplified by the whitefly *Bemisia tabaci*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **353**, 1757–1767.

Durand, R., Bouvresse, S., Berdjane, Z., Izri, A., Chosidow, O. & Clark, J.M. (2012) Insecticide resistance in head lice: clinical, parasitological and genetic aspects. *Clinical Microbiology and Infection*, **18**, 338–344.

Durden, L.A. & Musser, G.G. (1994) The sucking lice (Insecta, Anoplura) of the world: a taxonomic checklist with records of mammalian hosts and geographical distributions. *Bulletin of the American Museum of Natural History*, **218**, 50.

Escribá, M.C., Giardini, M.C. & Goday, C. (2011) Histone H3 phosphorylation and non-disjunction of the maternal X chromosome during male meiosis in sciarid flies. *Journal of Cell Science*, **124**, 1715–1725.

Esteban, M.R., Campos, M.C., Perondini, A.L. & Goday, C. (1997) Role of microtubules and microtubule organizing centers on meiotic chromosome elimination in *Sciara ocellaris*. *Journal of Cell Science*, **110**, 721–730.

de la Filia, A.G., Bain, S.A. & Ross, L. (2015) Haplodiploidy and the reproductive ecology of arthropods. *Current Opinion in Insect Science*, **9**, 36–43.

Gardner, A. & Ross, L. (2014) Mating ecology explains patterns of genome elimination. *Ecology Letters*, **17**, 1602–1612.

Goday, C. & Esteban, M.R. (2001) Chromosome elimination in sciarid flies. *BioEssays*, **23**, 242–250.

Goday, C. & Ruiz, M.F. (2002) Differential acetylation of histones H3 and H4 in paternal and maternal germline chromosomes during development of sciarid flies. *Journal of Cell Science*, **115**, 4765–4775.

Greciano, P.G. & Goday, C. (2006) Methylation of histone H3 at Lys4 differs between paternal and maternal chromosomes in *Sciara ocellaris* germline development. *Journal of Cell Science*, **119**, 4667–4677.

Havron, A., Rosen, D., Rössler, Y. & Hillel, J. (1987) Selection on the male hemizygous genotype in arrhenotokous insects and mites. *Entomophaga*, **32**, 261–268.

Hedrick, P.W. & Parker, J.D. (1997) Evolutionary genetics and genetic variation of haplodiploids and X-linked genes. *Annual Review of Ecology and Systematics*, **28**, 55–83.

Hindle, E. & Pontecorvo, G. (1942) Mitotic divisions following meiosis in *Pediculus corporis* males. *Nature*, **149**, 668.

Hodson, C.N., Hamilton, P.T., Dilworth, D., Nelson, C.J., Curtis, C.I. & Perlman, S.J. (2017) Paternal genome elimination in *Liposcelis* booklice (Insecta: Psocodea). *Genetics*, **206**, 1091–1100.

Hoffmann, A.A., Reynolds, K.T., Nash, M.A. & Weeks, A.R. (2008) A high incidence of parthenogenesis in agricultural pests. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 2473–2481.

Khosla, S., Mendiratta, G. & Brahmachari, V. (2006) Genomic imprinting in the mealybugs. *Cytogenetic and Genome Research*, **113**, 41–52.

Kuijper, B. & Pen, I. (2010) The evolution of haplodiploidy by male-killing endosymbionts: importance of population structure and endosymbiont mutualisms. *Journal of Evolutionary Biology*, **23**, 40–52.

Kwon, D.H., Kim, J.H., Kim, Y.H., Yoon, K.S., Clark, J.M. & Lee, S.H. (2014) Identification and characterization of an esterase involved in malathion resistance in the head louse *Pediculus humanus capitis*. *Pesticide Biochemistry and Physiology*, **112**, 13–18.

Leo, N.P., Campbell, N.J.H., Yang, X., Mumcuoglu, K. & Barker, S.C. (2002) Evidence from mitochondrial DNA that head lice and body lice of humans (Phthiraptera: Pediculidae) are conspecific. *Journal of Medical Entomology*, **39**, 662–666.

Li, H., Shao, R., Song, N. *et al.* (2015) Higher-level phylogeny of paraneopteran insects inferred from mitochondrial genome sequences. *Scientific Reports*, **5**, 8527.

McMeniman, C.J. & Barker, S.C. (2005) Transmission ratio distortion in the human body louse, *Pediculus humanus* (Insecta: Phthiraptera). *Heredity*, **96**, 63–68.

Meinking, T.L. (1999) Infestations. *Current Problems in Dermatology*, **11**, 73–118.

Normark, B.B. (2003) The evolution of alternative genetic systems in insects. *Annual Review of Entomology*, **48**, 397–423.

Normark, B.B. (2004) Haplodiploidy as an outcome of coevolution between male-killing cytoplasmic elements and their hosts. *Evolution*, **58**, 790–798.

Oakeshott, J.G., Horne, I., Sutherland, T.D. & Russell, R.J. (2003) The genomics of insecticide resistance. *Genome Biology*, **4**, 202.

Perotti, M.A., Catalá, S.S., del V Ormeño, A., Żelazowska, M., Biliński, S.M. & Braig, H.R. (2004) The sex ratio distortion in the human head louse is conserved over time. *BMC Genetics*, **5**, 10.

Prantera, G. & Bongiorni, S. (2012, 2012) Mealybug chromosome cycle as a paradigm of epigenetics. *Genetics Research International*, 867390, 1–11.

Raoult, D. & Roux, V. (1999) The body louse as a vector of reemerging human diseases. *Clinical Infectious Diseases*, **29**, 888–911.

Raymond, M. & Rousset, F. (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Ross, L., Shuker, D.M., Normark, B.B. & Pen, I. (2012) The role of endosymbionts in the evolution of haploid-male genetic systems in scale insects (Coccoidea). *Ecology and Evolution*, **2**, 1071–1081.

Ross, L., Hardy, N.B., Okusu, A. & Normark, B.B. (2013) Large population size predicts the distribution of asexuality in scale insects. *Evolution*, **67**, 196–206.

Rousset, F. (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Scholl, H. (1955) Ein Beitrag zur Kenntnis der Spermatogenese der Mallophagen. *Chromosoma*, **7**, 271–274.

Takano-Lee, M., Yoon, K.S., Edman, J.D., Mullens, B.A. & Clark, J.M. (2003) *In vivo* and *in vitro* rearing of *Pediculus humanus capitis* (Anoplura: Pediculidae). *Journal of Medical Entomology*, **40**, 628–635.

Tombesi, M.L. & Papeschi, A.G. (1993) Meiosis in *Haematopinus suis* and *Menacanthus stramineus* (Phthiraptera, Insecta). *Hereditas*, **119**, 31–38.

Tombesi, M.L., Papeschi, A.G. & Mola, L.M. (1999) Spermatogenesis in *Bovicola limbata* Gervais, 1844 and *B. caprae* Gurlt, 1843 (Phthiraptera, Ischnocera). *Cytologia*, **64**, 25–27.

Viera, A., Page, J. & Rufas, J. (2008) Inverted meiosis: the true bugs as a model to study. *Genome Dynamics*, **5**, 137–156.

Weir, B.S. & Cockerham, C.C. (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wright, S. (1933) Inbreeding and homozygosis. *Proceedings of the National Academy of Sciences of the United States of America*, **19**, 411–420.

Yoon, K.S., Strycharz, J.P., Gao, J.-R., Takano-Lee, M., Edman, J.D. & Clark, J.M. (2006) An improved *in vitro* rearing system for the human head louse allows the determination of resistance to formulated pediculicides. *Pesticide Biochemistry and Physiology*, **86**, 195–202.

Yoon, K.S., Symington, S.B., Lee, S.H., Soderlund, D.M. & Clark, J.M. (2008) Three mutations identified in the voltage-sensitive sodium channel α-subunit gene of permethrin-resistant human head lice reduce the permethrin sensitivity of house fly Vssc1 sodium channels expressed in Xenopus oocytes. *Insect Biochemistry and Molecular Biology*, **38**, 296–306.

Yoshizawa, K. & Johnson, K.P. (2010) How stable is the 'Polyphyly of Lice' hypothesis (Insecta: Psocodea)? A comparison of phylogenetic signal in multiple genes. *Molecular Phylogenetics and Evolution*, **55**, 939–951.

Accepted 2 November 2017
First published online 20 December 2017